Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT On Aug 19, 2003 this sequence version replaced gi:15530217.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 14 Row: m Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 34222363.

FEATURES Location/Qualifiers
source 1. .2096
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:11168 IMAGE:3842949"
/tissue_type="Placenta; choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
gene 1. .2096
/gene="BAIAP2L1"
/note="synonym: IRTKS"
/db_xref="GeneID:55971"
CDS 217. .1752
/gene="BAIAP2L1"
/codon_start=1
/product="BAI1-associated protein 2-like 1"
/protein_id="AAH13888.1"
/db_xref="GI:15530218"
/db_xref="GeneID:55971"
/translation="MSRGPEEVNRLTESTYRNVMEQFNPGLRNLINLGKNYEKAVNAM
ILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNESLDENFKKFHKEIIH
ELEKKIELDVKYMNATLKRYQTEHKNKLESLEKSQAELKKIRRKSQGSRNALKYEHKE
IEYVETVTSRQSEIQKFIADGCKEALLEEKRRFCFLVDKHCGFANHIHYYHLQSAELL
NSKLPRWQETCVDAIKVPEKIMNMIEEIKTPASTPVSGTPQASPMIERSNVVRKDYDT
LSKCSPKMPPAPSGRAYTSPLIDMFNNPATAAPNSQRVNNSTGTSEDPSLQRSVSVAT
GLNMMKKQKVKTIFPHTAGSNKTLLSFAQGDVITLLIPEEKDGWLYGEHDVSKARGWF
PSSYTKLLEENETEAVTVPTPSPTPVRSISTVNLSENSSVVIPPPDYLECLSMGAAAD
RRADSARTTSTFKAPASKPETAAPNDANGTAKPPFLSGENPFATVKLRPTVTNDRSAP
IIR"

ORIGIN

Query Match 100.0%; Score 2096; DB 8; Length 2096;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60
Db    1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60

Qy   61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120
Db   61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120

Qy  121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180
Db  121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180

Qy  181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240
Db  181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240

Qy  241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
Db  241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300

Qy  301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360
Db  301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360

Qy  361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
Db  361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Qy  421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
Db  421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480

Qy  481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540
Db  481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540

Qy  541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600
Db  541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600

Qy  601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660
Db  601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660

Qy  661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720
Db  661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720

Qy  721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780
Db  721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780

Qy  781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840
Db  781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840

Qy  841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900
Db  841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900

Qy  901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960
Db  901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960

Qy  961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020
Db  961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020

Qy 1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
Db 1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080

Qy 1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140
Db 1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140

Qy 1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200
Db 1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200

Qy 1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
Db 1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260

Qy 1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320
Db 1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320

Qy 1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380
```

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:39:55 ; Search time 10285 Seconds
(without alignments)
11584.237 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.....aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2096	100.0	2096	8	BC013888	BC013888 Homo sapi
	2	2025.2	96.6	2527	8	AF119666	AF119666 Homo sapi
	3	1821.4	86.9	2317	6	AX058219	AX058219 Sequence
	4	1560	74.4	1580	8	AK027109	AK027109 Homo sapi
	5	1235.4	58.9	1491	6	CQ730103	CQ730103 Sequence
	6	1163	55.5	3040	9	BC015459	BC015459 Mus muscu
	7	1042.8	49.8	2141	9	BC089216	BC089216 Rattus no
	8	859	41.0	2521	8	AK027142	AK027142 Homo sapi
	9	745.2	35.6	2064	5	BX933564	BX933564 Gallus ga
	10	671	32.0	2296	8	AK074419	AK074419 Homo sapi
	11	594.4	28.4	3616	5	BC061676	BC061676 Xenopus l
	12	556	26.5	559	6	AX379210	AX379210 Sequence
	13	397	18.9	685	5	BX931244	BX931244 Gallus ga
	14	382.6	18.3	2563	5	BC068330	BC068330 Danio rer
	15	381	18.2	2566	5	AY398378	AY398378 Danio rer
	16	377.8	18.0	2570	5	BC050238	BC050238 Danio rer
	17	365.8	17.5	2603	5	BC092786	BC092786 Danio rer
c	18	357	17.0	11840	8	AC093169	AC093169 Homo sapi
	19	355	16.9	3128	6	CQ843301	CQ843301 Sequence
	20	355	16.9	3128	8	AK124269	AK124269 Homo sapi
	21	307.2	14.7	1678	9	AF390179	AF390179 Mus muscu
	22	285.2	13.6	490	5	BX930239	BX930239 Gallus ga
	23	266.6	12.7	2078	9	CGU41899	U41899 Cricetinae
c	24	265.4	12.7	132072	8	AC004841	AC004841 Homo sapi
c	25	265.4	12.7	156195	8	AC093799	AC093799 Homo sapi
	26	265	12.6	1798	9	AY037934	AY037934 Rattus no
	27	265	12.6	2111	9	BC074009	BC074009 Rattus no
	28	261.6	12.5	1798	9	AF390178	AF390178 Mus muscu
	29	261.6	12.5	2545	9	BC006620	BC006620 Mus muscu
	30	261.6	12.5	2548	9	BC016411	BC016411 Mus muscu
	31	259.6	12.4	2080	6	AR069610	AR069610 Sequence
	32	259.6	12.4	2080	6	BD132800	BD132800 Insulin r
	33	259.6	12.4	2080	6	AR352563	AR352563 Sequence
	34	258	12.3	1078	8	HSU70669	U70669 Human Fas-l
	35	258	12.3	2033	8	AB017120	AB017120 Homo sapi
	36	258	12.3	2072	8	BC014020	BC014020 Homo sapi
	37	258	12.3	2077	8	AB169737	AB169737 Macaca fa
	38	258	12.3	2129	6	CQ896463	CQ896463 Sequence
	39	258	12.3	2129	8	AB015020	AB015020 Homo sapi
	40	258	12.3	2137	6	BD226257	BD226257 Pancreati
	41	258	12.3	2137	6	AX011646	AX011646 Sequence
	42	258	12.3	2519	8	BC032559	BC032559 Homo sapi
	43	258	12.3	2877	8	AB017119	AB017119 Homo sapi
	44	258	12.3	3168	8	AB015019	AB015019 Homo sapi
	45	256.4	12.2	1910	8	AK222670	AK222670 Homo sapi

ALIGNMENTS

RESULT 1
BC013888
LOCUS BC013888 2096 bp mRNA linear PRI 15-FEB-2005
DEFINITION Homo sapiens BAI1-associated protein 2-like 1, mRNA (cDNA clone MGC:11168 IMAGE:3842949), complete cds.
ACCESSION BC013888
VERSION BC013888.2 GI:33871229
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2096)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2096)
AUTHORS Director MGC Project.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

```
Db      1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380

Qy      1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Db      1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Qy      1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500

Db      1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500

Qy      1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560

Db      1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560

Qy      1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620

Db      1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620

Qy      1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680

Db      1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680

Qy      1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740

Db      1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740

Qy      1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800

Db      1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800

Qy      1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Db      1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Qy      1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920

Db      1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920

Qy      1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Db      1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Qy      1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAA 2040

Db      1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAA 2040

Qy      2041 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096

Db      2041 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
```

```
RESULT 2
AF119666
LOCUS       AF119666               2527 bp    mRNA    linear   PRI 26-AUG-2003
DEFINITION  Homo sapiens insulin receptor tyrosine kinase substrate mRNA,
            complete cds.
ACCESSION   AF119666
VERSION     AF119666.2  GI:34223696
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2527)
  AUTHORS   Peng,Y., Huang,C., Wu,T., Huang,Q., Gu,Y., Zhang,L., Li,Y., Xu,S.,
            Fu,G., Wang,Y., Chen,Z. and Han,Z.
  TITLE     A novel gene expressed in the human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2527)
  AUTHORS   Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
            Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-1999) Chinese National Human Genome Center at
            Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, P. R. China
REFERENCE   3  (bases 1 to 2527)
  AUTHORS   Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
            Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-AUG-2003) Chinese National Human Genome Center at
            Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, P. R. China
  REMARK    Sequence update by submitter
COMMENT     On Aug 26, 2003 this sequence version replaced gi:6563257.
FEATURES             Location/Qualifiers
     source          1. .2527
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
     CDS             217. .1752
                     /note="IRTKS"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="insulin receptor tyrosine kinase substrate"
                     /protein_id="AAF17223.2"
                     /db_xref="GI:34223697"
                     /translation="MSRGPEEVNRLTESTYRNVMEQFNPGLRNLINLGKNYEKAVNAM
                     ILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNESLDENFKKFHKEIIH
                     ELEKKIELDVKYMNATLKRYQTEHKNKLESLEKSQAELKKIRRKSQGSRNALKYEHKE
                     IEYVETVTSRQSEIQKFIADGCKEALLEEKRRFCFLVDKHCGFANHIHYYHLQSAELL
                     NSKLPRWQETCVDAIKVPEKIMNMIEEIKTPASTPVSGTPQASPMIERSNVVRKDYDT
                     LSKCSPKMPPAPSGRAYTSPLIDMFNNPATAAPNSQRVNNSTGTSEDPSLQRSVSVAT
                     GLNMMKKQKVKTIFPHTAGSNKTLLSFAQGDVITLLIPEEKDGWLYGEHDVSKARGWF
                     PSSYTKLLEENETEAVTVPTPSPTPVRSISTVNLSENSSVVIPPPDYLECLSMGAAAD
                     RRADSARTTSTFKAPASKPETAAPNDANGTAKPPFLSGENPFATVKLRPTVTNDRSAP
                     IIR"
ORIGIN
```

```
  Query Match           96.6%;  Score 2025.2;  DB 8;  Length 2527;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 2027;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60

Db          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60

Qy         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120

Db         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120

Qy        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180

Db        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180

Qy        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240

Db        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240

Qy        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300

Db        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300

Qy        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360

Db        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360

Qy        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Db        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Qy        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480

Db        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
```

```
Qy        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540
Db        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540

Qy        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600
Db        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600

Qy        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660
Db        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660

Qy        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720
Db        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720

Qy        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780
Db        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780

Qy        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840
Db        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840

Qy        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900
Db        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900

Qy        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960
Db        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960

Qy        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020
Db        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020

Qy       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
Db       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080

Qy       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140
Db       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140

Qy       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200
Db       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200

Qy       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
Db       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260

Qy       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320
Db       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320

Qy       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380
Db       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380

Qy       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440
Db       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Qy       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500
Db       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500

Qy       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560
Db       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCCTATCCATGGGGGCAGCT 1560

Qy       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620
Db       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620

Qy       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680
Db       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680

Qy       1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740
Db       1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCAGCACCC 1740

Qy       1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800
Db       1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800

Qy       1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860
Db       1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Qy       1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920
Db       1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920

Qy       1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980
Db       1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Qy       1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
Db       1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
```

```
RESULT 3
AX058219
LOCUS       AX058219               2317 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 89 from Patent WO0077040.
ACCESSION   AX058219
VERSION     AX058219.1  GI:12310720
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Yang,J., Reddy,R. and Lu,D.A.
  TITLE     Human intracellular signaling molecules
  JOURNAL   Patent: WO 0077040-A 89 21-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2317
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           86.9%;  Score 1821.4;  DB 6;  Length 2317;
  Best Local Similarity   99.9%;  Pred. No. 6.2e-306;
  Matches 1822;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        208 GCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGG 267
Db         30 GCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGG 89

Qy        268 AATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTAT 327
Db         90 AATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTAT 149

Qy        328 GAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAG 387
Db        150 GAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAG 209
```

```
Qy       388 ATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAG 447
Db       210 ATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAG 269

Qy       448 ATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 507
Db       270 ATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 329

Qy       508 AAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCA 567
Db       330 AAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCA 389

Qy       568 ACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAA 627
Db       390 ACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAA 449

Qy       628 GCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAA 687
Db       450 GCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAA 509

Qy       688 CACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 747
Db       510 CACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 569

Qy       748 ATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTT 807
Db       570 ATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTT 629

Qy       808 GATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTA 867
Db       630 GATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTA 689

Qy       868 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAG 927
Db       690 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAG 749

Qy       928 AAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACT 987
Db       750 AAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACT 809

Qy       988 CCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTT 1047
Db       810 CCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTT 869

Qy      1048 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTG 1107
Db       870 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTG 929

Qy      1108 ATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCA 1167
Db       930 ATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCA 989

Qy      1168 ACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAAC 1227
Db       990 ACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAAC 1049

Qy      1228 ATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACC 1287
Db      1050 ATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACC 1109

Qy      1288 TTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1347
Db      1110 TTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1169

Qy      1348 TGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACG 1407
Db      1170 TGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACG 1229

Qy      1408 AAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCA 1467
Db      1230 AAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCA 1289

Qy      1468 GTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCC 1527
Db      1290 GTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCC 1349

Qy      1528 GACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGG 1587
Db      1350 GACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGG 1409

Qy      1588 ACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCC 1647
Db      1410 ACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCC 1469

Qy      1648 AACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTC 1707
Db      1470 AACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTC 1529

Qy      1708 CGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGA 1767
Db      1530 CGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGA 1589

Qy      1768 CTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGT 1827
Db      1590 CTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGT 1649

Qy      1828 GCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAA 1887
Db      1650 GCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAA 1709

Qy      1888 TCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGT 1947
Db      1710 TCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGT 1769

Qy      1948 AATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAAC 2007
Db      1770 AATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAAC 1829

Qy      2008 AGTAAACTTGGTTTCAATCTTTA 2030
Db      1830 AGTAAACTTGGTTTCAATCTTTA 1852
```

RESULT 4
AK027109

LOCUS	AK027109 1580 bp mRNA linear PRI 12-SEP-2003
DEFINITION	Homo sapiens cDNA: FLJ23456 fis, clone HSI07211, highly similar to AF119666 Homo sapiens insulin receptor tyrosine kinase substrate mRNA.
ACCESSION	AK027109
VERSION	AK027109.1 GI:10440150
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1580)
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE	Direct Submission
JOURNAL	Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of

International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Departent of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

```
FEATURES             Location/Qualifiers
     source          1. .1580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07211"
                     /tissue_type="human small intestine"
                     /clone_lib="HSI"
                     /note="cloning vector pME18SFL3"
     misc_feature    1. .1580
                     /note="highly similar to AF119666 Homo sapiens insulin
                     receptor tyrosine kinase substrate mRNA"
ORIGIN
```

Query Match 74.4%; Score 1560; DB 8; Length 1580;
Best Local Similarity 99.6%; Pred. No. 1.4e-260;
Matches 1574; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy        499 AAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATAT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATAT 60

Qy        559 ATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAG 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAG 120

Qy        619 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTC 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTC 180

Qy        679 AAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATC 240

Qy        739 CAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 300

Qy        799 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCT 360

Qy        859 GCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAA 420

Qy        919 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTG 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTG 480

Qy        979 TCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTAC 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTAC 540

Qy       1039 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACC 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACC 600

Qy       1099 AGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 660

Qy       1159 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACG 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACG 720

Qy       1219 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCC 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCC 780

Qy       1279 AACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAG 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAG 840

Qy       1339 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCG 1398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCG 900

Qy       1399 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGC 1458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGC 960

Qy       1459 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1020

Qy       1519 CCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGAT 1578
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCCCACCCGACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGAT 1080

Qy       1579 TCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCT 1638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCT 1140

Qy       1639 AACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACT 1698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACT 1200

Qy       1699 GTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAC 1758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAC 1260

Qy       1759 AGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGT 1818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGT 1320

Qy       1819 CTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGC 1878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGC 1380

Qy       1879 TTAAGCAAATCATG-CTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1937
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTAAGCAAATCATGCCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1440

Qy       1938 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1500

Qy       1998 TTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAA 2057
              |||||||||||||||||||||||||||||||    ||||||||||||||||||||||||
Db       1501 TTCAAGAAACAGTAAACTTGGTTTCAATCTTTACTGTAAAAAAAAAAAAAAAAAAAAAA 1560

Qy       2058 AAAAAAAAAAAAAAAAAAAA 2077
              ||||||||||||||||||||
Db       1561 AAAAAAAAAAAAAAAAAAAA 1580
```

```
RESULT 5
CQ730103
LOCUS       CQ730103               1491 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 16037 from Patent WO02068579.
ACCESSION   CQ730103
VERSION     CQ730103.1  GI:42303049
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 16037 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1491
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           58.9%;  Score 1235.4;  DB 6;  Length 1491;
  Best Local Similarity  88.8%;  Pred. No. 3e-204;
  Matches 1436;  Conservative    0;  Mismatches    1;  Indels  180;  Gaps    2;

Qy         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 60

Qy        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 120

Qy        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 180

Qy        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 240

Qy        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 300

Qy        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 360

Qy        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 420

Qy        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 480

Qy        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600
               ||||||||||||||||||||||
Db        481 GAACTTGACGTGAAATATATGAAC---------------------------------- 504

Qy        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660
Db        505 ---------------------------------------------------------- 504

Qy        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720
                                                   ||||||||||||||||||
Db        505 ----------------------------------------TATGTGGAGACCGTTACT 522

Qy        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 582

Qy        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 642

Qy        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 702

Qy        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 762

Qy        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 822

Qy       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
               ||                                          ||||||||||||||||
Db        823 GT------------------------------------------GATGCCCCCCGCTCCT 840

Qy       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 900

Qy       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 960

Qy       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1020

Qy       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1080

Qy       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1140

Qy       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1200

Qy       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1260

Qy       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1261 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTATCCATGGGGGCAGCT 1320

Qy       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1380

Qy       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGC 1677
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGC 1437

RESULT 6
BC015459
LOCUS       BC015459                3040 bp    mRNA    linear   ROD 01-MAR-2005
DEFINITION  Mus musculus BAI1-associated protein 2-like 1, mRNA (cDNA clone
            MGC:8269 IMAGE:3592632), complete cds.
ACCESSION   BC015459
VERSION     BC015459.1  GI:15930030
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3040)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
```

Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 3040)

AUTHORS Director MGC Project.

TITLE Direct Submission

JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 11 Row: b Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229018.

FEATURES Location/Qualifiers

source 1. .3040
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:8269 IMAGE:3592632"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

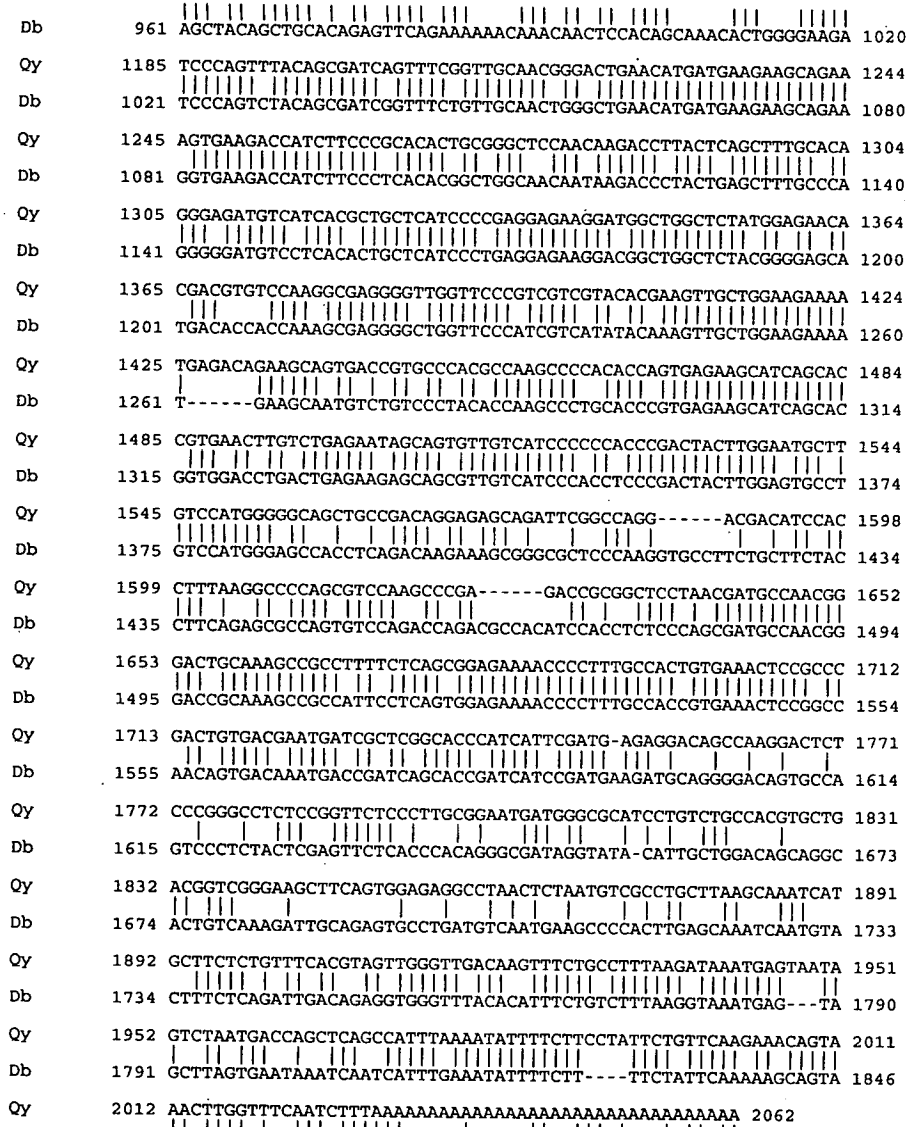
gene 1. .3040
/gene="Baiap2l1"
/db_xref="GeneID:66898"
/db_xref="MGI:1914148"

CDS 50. .1594
/gene="Baiap2l1"
/codon_start=1
/product="BAI1-associated protein 2-like 1"
/protein_id="AAH15459.1"
/db_xref="GI:15930031"
/db_xref="GeneID:66898"
/db_xref="MGI:1914148"
/translation="MSRGPEEVNRLTENTYRNVMEQFNPGLRNLINLGKNYEKAVNAM
ILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNETLDENFKKFHKDIIH
ELEKKTELDVKYMNATLKRYQAEHRNKLDSLEKSQAELKKIRRKSQGGRNALKYEHKE
IEYVETVTSRQSEIQKFIADGCKEALLEEKRRFCFLVDKHCSFASHIHYYHMQSAELL
NSKLPRWQETCCDATKVPEKIMNMIEEIKTPISTPVSGTPQPSPMIERSKMIGKDYDT
LSKYSPKMPPAPSVKAYTSPLIDMFNNPATAAQSSEKTNNSTANTGEDPSLQRSVSVA
TGLNMMKKQKVKTIFPHTAGNNKTLLSFAQGDVLTLLIPEEKDGWLYGEHDTTKARGW
FPSSYTKLLEENEAMSVPTPSPAPVRSISTVDLTEKSSVVIPPPDYLECLSMGATSDK
KAGAPKVPSASTFRAPVSRPDATSTSPSDANGTAKPPFLSGENPFATVKLRPTVTNDR
SAPIIR"

ORIGIN

Query Match 55.5%; Score 1163; DB 9; Length 3040;
Best Local Similarity 78.5%; Pred. No. 1.2e-191;
Matches 1501; Conservative 0; Mismatches 380; Indels 30; Gaps 8;

```
Qy     168 CGAGGCCCAGCCGCGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGG 227
Db       1 CGAGGCCCCGCCGCGTCCTCCCGCGCTCGCCCTCCCGGAGGCTGCAGCCATGTCTCGGGG  60

Qy     228 GCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAA 287
Db      61 GCCGGAGGAGGTGAACCGGCTCACGGAGAACACGTACCGGAATGTGATGGAACAATTCAA 120

Qy     288 TCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTAT 347
Db     121 CCCTGGACTACGGAACTTAATAAACCTAGGAAAAAATTATGAAAAAGCGGTGAATGCTAT 180

Qy     348 GATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGG 407
Db     181 GATCCTGGCAGGAAAAGCCTACTACGACGGCGTGGCCAAGATTGGTGAAATTGCCACTGG 240

Qy     408 GTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAA 467
Db     241 GTCTCCTGTGTCAACAGAGCTAGGCCATGTCCTCATAGAGATTTCCAGCACCCACAAGAA 300

Qy     468 ACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCT 527
Db     301 ACTCAATGAGACTCTTGATGAAAATTTCAAAAAATTCCATAAGGATATTATCCATGAGCT 360

Qy     528 GGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAAC 587
Db     361 GGAGAAGAAGACAGAACTTGATGTAAAGTATATGAACGCTACTCTTAAAAGATATCAAGC 420

Qy     588 AGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAG 647
Db     421 AGAGCACAGGAATAAATTAGATTCTTTGGAGAAATCTCAGGCGGAGCTGAAGAAGATCAG 480

Qy     648 AAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGT 707
Db     481 AAGAAAAAGTCAAGGTGGACGAAACGCACTCAAATACGAGCACAAAGAAATTGAGTATGT 540

Qy     708 GGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGA 767
Db     541 GGAGACGGTGACGTCTCGCCAGAGTGAGATCCAGAAGTTCATTGCCGATGGCTGCAAAGA 600

Qy     768 GGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 827
Db     601 AGCCCTGCTTGAGGAAAAAAGGCGCTTCTGCTTTCTGGTGGACAAGCACTGCAGCTTTGC 660

Qy     828 AAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCG 887
Db     661 CAGTCACATACACTACTATCACATGCAGTCTGCAGAGTTACTTAATTCCAAGCTGCCCCG 720

Qy     888 GTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGA 947
Db     721 GTGGCAGGAAACCTGTTGTGATGCCACAAAAGTCCCAGAGAAAATCATGAACATGATAGA 780

Qy     948 AGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGAT 1007
Db     781 AGAGATAAAGACGCCCATCTCCACCCCCGTGTCCGGGACGCCACAGCCATCGCCCATGAT 840

Qy    1008 CGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGAT 1067
Db     841 CGAGAGAAGCAAGATGATTGGGAAAGATTATGACACCCTTTCTAAATACTCACCAAAGAT 900

Qy    1068 GCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCC 1127
Db     901 GCCCCCTGCTCCTTCAGTCAAAGCCTACACCAGTCCTTTGATTGACATGTTCAATAACCC 960

Qy    1128 AGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAG---GTACTTCCGAAGA 1184
```



```
Db        961 AGCTACAGCTGCACAGAGTTCAGAAAAAACAAACAACTCCACAGCAAACACTGGGGAAGA 1020
Qy       1185 TCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAA 1244
Db       1021 TCCCAGTCTACAGCGATCGGTTTCTGTTGCAACTGGGCTGAACATGATGAAGAAGCAGAA 1080
Qy       1245 AGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACA 1304
Db       1081 GGTGAAGACCATCTTCCCTCACACGGCTGGCAACAATAAGACCCTACTGAGCTTTGCCCA 1140
Qy       1305 GGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACA 1364
Db       1141 GGGGGATGTCCTCACACTGCTCATCCCTGAGGAGAAGGACGGCTGGCTCTACGGGGAGCA 1200
Qy       1365 CGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAA 1424
Db       1201 TGACACCACCAAAGCGAGGGGCTGGTTCCCATCGTCATATACAAAGTTGCTGGAAGAAAA 1260
Qy       1425 TGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCAC 1484
Db       1261 T------GAAGCAATGTCTGTCCCTACACCAAGCCCTGCACCCGTGAGAAGCATCAGCAC 1314
Qy       1485 CGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTT 1544
Db       1315 GGTGGACCTGACTGAGAAGAGCAGCGTTGTCATCCCACCTCCCGACTACTTGGAGTGCCT 1374
Qy       1545 GTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGG------ACGACATCCAC 1598
Db       1375 GTCCATGGGAGCCACCTCAGACAAGAAAGCGGGCGCTCCCAAGGTGCCTTCTGCTTCTAC 1434
Qy       1599 CTTTAAGGCCCCAGCGTCCAAGCCCGA------GACCGCGGCTCCTAACGATGCCAACGG 1652
Db       1435 CTTCAGAGCGCCAGTGTCCAGACCAGACGCCACATCCACCTCTCCCAGCGATGCCAACGG 1494
Qy       1653 GACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCC 1712
Db       1495 GACCGCAAAGCCGCCATTCCTCAGTGGAGAAAACCCCTTTGCCACCGTGAAACTCCGGCC 1554
Qy       1713 GACTGTGACGAATGATCGCTCGGCACCCATCATTCGATG-AGAGGACAGCCAAGGACTCT 1771
Db       1555 AACAGTGACAAATGACCGATCAGCACCGATCATCCGATGAAGATGCAGGGGACAGTGCCA 1614
Qy       1772 CCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGCTG 1831
Db       1615 GTCCCTCTACTCGAGTTCTCACCCACAGGGCGATAGGTATA-CATTGCTGGACAGCAGGC 1673
Qy       1832 ACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATCAT 1891
Db       1674 ACTGTCAAAGATTGCAGAGTGCCTGATGTCAATGAAGCCCCACTTGAGCAAATCAATGTA 1733
Qy       1892 GCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATA 1951
Db       1734 CTTTCTCAGATTGACAGAGGTGGGTTTACACATTTCTGTCTTTAAGGTAAATGAG---TA 1790
Qy       1952 GTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTA 2011
Db       1791 GCTTAGTGAATAAATCAATCATTTGAAATATTTTCTT----TTCTATTCAAAAAGCAGTA 1846
Qy       2012 AACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2062
Db       1847 AATTTGGCTCAAATTTTTAAATTTTATTTTTAAGGAAATAGCCAGAATAA 1897
```

```
RESULT 7
BC089216
LOCUS       BC089216               2141 bp    mRNA    linear   ROD 30-MAR-2005
DEFINITION  Rattus norvegicus similar to RIKEN cDNA 1300006M19 (predicted),
            mRNA (cDNA clone IMAGE:7303117), partial cds.
ACCESSION   BC089216
VERSION     BC089216.1  GI:58402628
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 2141)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 2141)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
            College of Wisconsin
            cDNA Library Preparation: Express Genomics
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 202 Row: l Column: 24
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES              Location/Qualifiers
     source           1. .2141
                      /organism="Rattus norvegicus"
                      /mol_type="mRNA"
                      /db_xref="taxon:10116"
                      /clone="IMAGE:7303117"
                      /tissue_type="Ovary, rat"
                      /clone_lib="NIH_MGC_253"
                      /lab_host="DH10B"
                      /note="Vector: pExpress1"
     gene             <1. .2141
                      /gene="RGD1308452_predicted"
```

```
                     /db_xref="GeneID:304282"
                     /db_xref="RGD:1308452"
     CDS             <1..1620
                     /gene="RGD1308452_predicted"
                     /codon_start=1
                     /product="RGD1308452_predicted protein"
                     /protein_id="AAH89216.1"
                     /db_xref="GI:58402629"
                     /db_xref="GeneID:304282"
                     /db_xref="RGD:1308452"
                     /translation="WFSPADSCGGSRLPRAAQSQALGSPAARSPEATRPRRVLPRSPS
                     RRLHHVSGAGGGEPAHGEHVPAMILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEIS
                     STHKKLNETLDENFKKFHKEIIHELEKKTELDVKYMNATLKRYQAEHRNKLDSLEKSQ
                     AELKKIRRKSQGGRNALKYEHKEIEYVETVTSRQSEIQKFIADGCKEALLEEKRRFCF
                     LVDKHCSFASHIHRYHLQSAELLNSKLPRWQETCCDATKVPEKIMNMIEEIKTPISTP
                     VSGTPQPSPMTERSKMIGKDYDTLSKYSPKMPPAPSVKAYTSPLIDMFNNPATAGQSA
                     EKTNNSTANTGDDPSLQRSVSVATGLNMMKKQKVKTIFPHTAGNNKTLLSFAQGDVLT
                     LLIPEEKDGWLYGEHDTTKVRGWFPSSYTKLLEENMKEAMSVPTPSSAPVRSISTVDL
                     TEKSSVVIPPPDYLECLSMGATSDKRADAAKIPSTSTFKAPVPRPDATSTSPSDSNGT
                     AKPPFLSGENPFATVKLRPTVTNDRSAPIIR"
ORIGIN

  Query Match             49.8%;  Score 1042.8;  DB 9;  Length 2141;
  Best Local Similarity   75.8%;  Pred. No. 8.3e-171;
  Matches 1466;  Conservative    0;  Mismatches  362;  Indels  106;  Gaps   10;

Qy         75 GGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGCCCCTTAGTCGCGCC 134
Db          2 GGTTCTCTCCGGCTGACAGCTGTGGCGGCAGCCGACTCCCGAGGGCCGCTCAGTCGCAAG 61

Qy        135 CC---GGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCGCGTCCTCCCGC 191
Db         62 CCCTAGGTTCTCCCGCCGCCCGCAGTCCGGAGGCCACGAGGCCCCGCCGCGTCCTCCCGC 121

Qy        192 GCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCAC 251
Db        122 GCTCGCCCTCCCGGAGGCTGCA-CCATGTCTCGGGGGCCGGAGGAGGTGAACCGGCTCAC 180

Qy        252 GGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAA 311
Db        181 GGAGAACACGTACCG---------------------------------------------- 195

Qy        312 CCTGGGGAAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTA 371
Db        196 ------------------------------GCTATGATCCTGGCAGGAAAAGCCTACTA 224

Qy        372 CGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGG 431
Db        225 CGACGGAGTGGCCAAGATTGGGGAAATCGCCACTGGGTCTCCTGTGTCAACAGAACTAGG 284

Qy        432 ACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAA 491
Db        285 CCATGTTCTCATAGAGATTTCCAGCACCCACAAGAAACTCAACGAGACTCTCGATGAGAA 344

Qy        492 TTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGT 551
Db        345 TTTCAAAAAATTCCATAAGGAGATTATCCATGAGTTGGAGAAGAAGACAGAACTTGATGT 404

Qy        552 GAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTC 611
Db        405 AAAGTACATGAACGCTACTCTTAAAAGATATCAAGCAGAACACAGGAATAAATTAGATTC 464

Qy        612 TTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAA 671
Db        465 TTTGGAGAAATCTCAGGCGGAGTTGAAGAAGATCAGAAGAAAAAGTCAAGGTGGACGAAA 524

Qy        672 CGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAG 731
Db        525 CGCACTCAAATATGAACACAAAGAAATCGAGTATGTGGAGACAGTTACCTCTCGCCAGAG 584

Qy        732 TGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCG 791
Db        585 TGAGATCCAGAAGTTCATCGCAGATGGCTGCAAAGAAGCGCTGCTTGAGGAGAAGAGGCG 644

Qy        792 CTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTT 851
Db        645 CTTCTGCTTTCTGGTGGACAAGCACTGCAGCTTTGCCAGTCACATACACCGCTATCACCT 704

Qy        852 ACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGC 911
Db        705 GCAGTCTGCAGAATTACTTAACTCCAAGCTGCCCCGATGGCAGGAAACCTGTTGTGATGC 764

Qy        912 CATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTAC 971
Db        765 CACGAAAGTCCCAGAGAAAATCATGAACATGATAGAAGAGATAAAGACGCCCATCTCTAC 824

Qy        972 CCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAA 1031
Db        825 CCCAGTGTCCGGGACGCCACAGCCATCGCCCATGACGGAGAGAAGCAAGATGATTGGGAA 884

Qy       1032 AGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGC 1091
Db        885 AGACTACGACACCCTTTCTAAATACTCACCGAAGATGCCCCCCGCTCCTTCAGTCAAAGC 944

Qy       1092 ATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACA 1151
Db        945 CTACACCAGTCCTCTGATTGACATGTTCAACAACCCTGCCACGGCTGGACAGAGCGCAGA 1004

Qy       1152 AAGGGTAAATAATTCAACAG---GTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTC 1208
Db       1005 GAAAACCAACAACTCCACAGCAAACACTGGGGACGATCCCAGTTTACAGCGGTCGGTCTC 1064

Qy       1209 GGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACAC 1268
Db       1065 TGTTGCAACTGGGCTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCTCACAC 1124

Qy       1269 TGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCAT 1328
Db       1125 CGCTGGCAACAATAAGACCTTACTGAGCTTTGCCCAGGGGGACGTCCTCACACTGCTCAT 1184

Qy       1329 CCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTG 1388
Db       1185 TCCCGAGGAGAAGGACGGCTGGCTCTACGGGGAGCATGACACCACCAAAGTGAGGGGCTG 1244

Qy       1389 GTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCC 1448
Db       1245 GTTCCCATCATCATACACAAAGTTGCTGGAAGAAAACATGAAGGAAGCTATGTCTGTACC 1304

Qy       1449 CACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAG 1508
Db       1305 AACACCAAGCTCTGCACCAGTGAGAAGCATCAGCACTGTGGATCTGACTGAGAAGAGCAG 1364

Qy       1509 TGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAG 1568
Db       1365 CGTTGTCATCCCTCCTCCTGACTACCTGGAGTGCCTGTCCATGGGAGCCACCTCCGACAA 1424

Qy       1569 GAGAGCAGATTCGGCCAGGA------CGACATCCACCTTTAAGGCCCCAGCGTCCAAGCC 1622
Db       1425 GAGAGCAGACGCTGCCAAGATACCTTCTACTTCTACCTTCAAAGCCCCAGTGCCCAGACC 1484

Qy       1623 CGAGACCGCGG------CTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAG 1676
Db       1485 AGATGCCACGTCCACATCTCCCAGTGACTCAAATGGGACCGCGAAGCCACCATTCCTCAG 1544

Qy       1677 CGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGC 1736
Db       1545 TGGAGAAAACCCCTTTGCCACCGTGAAACTCCGGCCGACAGTGACCAATGACCGATCGGC 1604

Qy       1737 ACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTG 1796
Db       1605 ACCGATCATCCGGTGAAAATGCAG---AGGACCCTGCCGGAGTGGTTC---TCTCACCTG 1658

Qy       1797 CGGAATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAG 1856
Db       1659 CAGGGCGATAGGTCTTCACTGCTGGACAAGTGCAGGCACTGACAAAGCCTGCAGAGGGTG 1718
```

```
Qy       1857 GCCTAACTCTAATGTCGCCTGCTTAAGCAAATCA--TGCTTCTCTGTTTCACGTAGTTGG 1914
             |  | ||| |    | ||  ||| |||||||||      || ||||||| |   || |||
Db       1719 CCTGATCTCAATGAACCCCGACTTGAGCAAATCAATGTATTTTCTGTTTGATAGAGCTGG 1778

Qy       1915 GTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATT 1974
             | | |||    || || ||||||| ||||||||  |||  || ||    |  | |  | ||
Db       1779 GCTTACACAATTTTGTCTTTAAGGTAAATGAG---TAGCTTAGTGCATAAATGAATCGTT 1835

Qy       1975 TAAAATATTTTCTT 1988
             | || | |||||||
Db       1836 TGAAGTGTTTTCTT 1849
```

RESULT 8
AK027142
LOCUS AK027142 2521 bp mRNA linear PRI 12-SEP-2003
DEFINITION Homo sapiens cDNA: FLJ23489 fis, clone LNG00746.
ACCESSION AK027142
VERSION AK027142.1 GI:10440194
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2521)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Departent of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES Location/Qualifiers
source 1. .2521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LNG00746"
/tissue_type="human lung"
/clone_lib="LNG"
/note="cloning vector pME18SFL3"
CDS 65. .589
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15671.1"
/db_xref="GI:10440195"
/translation="MMKKQKVKTIFPHTAGSNKTLLSFAQGDVITLLIPEEKDGWLYG EHDVSKARGWFPSSYTKLLEENETEAVTVPTPSPTPVRSISTVNLSENSSVVIPPPDY LECLSMGAAADRRADSVRTTFTFKVPVFKPETAAPNDANGTAKPPFLSGENPFATVKL RPTVTNDRSAPIIR"
ORIGIN

Query Match 41.0%; Score 859; DB 8; Length 2521;
Best Local Similarity 99.4%; Pred. No. 6.6e-139;
Matches 862; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1164 TTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACT 1223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACT 60

Qy       1224 GAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAA 1283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAA 120

Qy       1284 GACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGA 1343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGA 180

Qy       1344 TGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTA 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTA 240

Qy       1404 CACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCAC 1463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCAC 300

Qy       1464 ACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCC 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCC 360

Qy       1524 ACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGC 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGT 420

Qy       1584 CAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGA 1643
             |||||||||||| |||||||||||| ||||| || ||||||||||||||||||||||||||||
Db        421 CAGGACGACATTCACCTTTAAGGTCCCAGTGTTCAAGCCCGAGACCGCGGCTCCTAACGA 480

Qy       1644 TGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAA 1703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAA 540

Qy       1704 ACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCA 1763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCA 600

Qy       1764 AGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCC 1823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCC 660

Qy       1824 ACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAG 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAG 720

Qy       1884 CAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAAT 1943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAAT 780

Qy       1944 GAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAG 2003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAG 840

Qy       2004 AAACAGTAAACTTGGTTTCAATCTTTA 2030
             ||||||||||||||||||||||||||||
Db        841 AAACAGTAAACTTGGTTTCAATCTTTA 867
```

RESULT 9
BX933564
LOCUS BX933564 2064 bp mRNA linear VRT 30-MAR-2004
DEFINITION Gallus gallus finished cDNA, clone ChEST46b24.
ACCESSION BX933564
VERSION BX933564.2 GI:46017940
KEYWORDS .
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

```
REFERENCE   1  (bases 1 to 2064)
  AUTHORS   Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
            Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
            Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
            Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
            Tickle,C. and Wilson,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT     On Apr 1, 2004 this sequence version replaced gi:41634092.
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
            sequencing project.
            This sequence is from the
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
            from a library constructed by Elizabeth Bosch. cDNA was prepared
            from RNA extracted from whole embryo, normalised, and poly
            A-trimmed. EcoRI-NotI cut cDNA was then ligated into the vector.
            Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host:
            Escherichia coli DH10B.
FEATURES             Location/Qualifiers
     source          1. .2064
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST46b24"
                     /clone_lib="CSEQCHN04"
                     /dev_stage="stage 10"
ORIGIN

  Query Match             35.6%;  Score 745.2;  DB 5;  Length 2064;
  Best Local Similarity   69.8%;  Pred. No. 3.7e-119;
  Matches 1088;  Conservative    0;  Mismatches  443;  Indels   27;  Gaps    5;

Qy        206 CGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACC 265
Db         30 CCGCCGGCACCATGTCGCGGGACCCCGAGGAGGTGAACAAGCTAACGGAGAGCACTTACA 89

Qy        266 GGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATT 325
Db         90 AAAATGTTATGGAGCAATTCAATCCAGGACTGCGGAATTTAATAAATCTTGGGAAGAATT 149

Qy        326 ATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCA 385
Db        150 ACGAAAAAGCTGTTAATGCTATGGTAGTCGCTGGAAGATCATACTACGACAGTCTGGCAA 209

Qy        386 AGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAG 445
Db        210 AGATTGGGGATATATCAGCAGATTCACCTGTTTCTAAAGAGTTAGGCCAAGTCCTGGTGG 269

Qy        446 AGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCC 505
Db        270 AAATTTCAAGAACACACAAGAAACTTAACGACAGCCTGGAGGAAAGTTTTAAGAAATTTC 329

Qy        506 ACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACG 565
Db        330 ATAAGGAAATTATATCTGAACTGGAGAAGAAAACAGACCTGGATGTAAAATACATGACCG 389

Qy        566 CAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCC 625
Db        390 CGACTTTAAAGAGGTACCAAACCGAACACAGAAGTAAATTGGATTCCCTGGAGAAGTCTC 449

Qy        626 AAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATG 685
Db        450 AGGCTGAGCTGAAAAAAATCCGAAGGAAAAGCCAAGGAGCACGAAATGTAATGAAATACG 509

Qy        686 AACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAAT 745
Db        510 AGCATAAGGAAATGGAGTATTTGGAAACCGTGAGCTCTCGACAGACTGACATTCAGAGAT 569

Qy        746 TCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGG 805
Db        570 TTATTGCTGAAGGCTGTAGAGAAGCTCTACTTGAAGAAAAAAGAAGATTCTGTTTTCTAG 629

Qy        806 TTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAAC 865
Db        630 TTGACAAGCACTGCAGTTTTACCCAGCACATGCACTTCTATCACGCTCAGTGTGCAGAAT 689

Qy        866 TACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAG 925
Db        690 TCCTCAACGCCAAGCTGCCTGGGTGGCAGGAAATCTGCAGTGATGCCACCAAAGTTCCTG 749

Qy        926 AGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAA 985
Db        750 AAAAAGTCAAAATGATGATTGAGGAAATAAGGACACCTGGTTCCACTCCCATATCAGGAA 809

Qy        986 CTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCC 1045
Db        810 CTCCGCAGCCTTCACCAATGATAGAGAGAAACACCATGATTGGAAACAATTTTTATGCTC 869

Qy       1046 TTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCT 1105
Db        870 ATAATGAAAATGCATCAAAGGTGCCCCCTGCTCCTGTGGGCAGGGCATACACCAGCCCGC 929

Qy       1106 TGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCAC---AAAGGGTAAATA 1162
Db        930 TTGTGGACATGTTCAACAATCCTGCCATGGCTCCAAAGGTGTCTCCTGAAAGACCGACCA 989

Qy       1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 1222
Db        990 ACCCAGCAGAGAATTCAGATGATGCCAGTTTATCACGCTCAACTTCTGTTGCCACAGGGC 1049

Qy       1223 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 1282
Db       1050 TAAATATAATGAAAAGGCAGAAAGTAAGGACAATCTTTCCACATACTGCTGGAAATAACA 1109

Qy       1283 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 1342
Db       1110 AGACATTACTTAGCTTTGCACAGGGAGATACCATCATACTTCTTGTATCTGAAGAAAAGG 1169

Qy       1343 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 1402
Db       1170 ATGGCTGGCTTTACGGGGAACATGAGTCAACTAAAGTAAAAGGATGGTTTCCATCATCAT 1229

Qy       1403 ACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCA 1462
Db       1230 ATACCAGACCACTAGAAGAATCAGA---AGAAAAAGCCTTTGTTCCAGCATCAAGCGCTG 1286

Qy       1463 CACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCC 1522
Db       1287 CGCCGGTTAGAAGTATTAGCTCAGTGAATCTTGCAGAAAAAGCTGGTGTTGTCCTCCCAC 1346

Qy       1523 CACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGG 1582
Db       1347 CGCCAGACTATCTGGGTT----------------CAACTTCAGATAAGAGAATGGAGTCTT 1391

Qy       1583 CCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAA-- 1640
Db       1392 CCAAAGGTACTGCTGC---TAGAACACCAACTGACAAACCAGAAAGTGCAGGTCTGAAAC 1448

Qy       1641 -CGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTG 1699
Db       1449 CCGACATGAATGGCATCGTAAAACCACCCTTTCTAAGTGGAGAAAATCCTTTTGCAACTG 1508

Qy       1700 TGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGA 1757
Db       1509 TGAAACTCAGACCAACAGTAACAAATGACAGATCAGCACCCATTATTCGATGAACAGA 1566

RESULT 10
AK074419
LOCUS       AK074419               2296 bp    mRNA    linear   PRI 12-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ23839 fis, clone KAT02564.
ACCESSION   AK074419
VERSION     AK074419.1  GI:18677016
KEYWORDS    oligo capping; fis (full insert sequence).
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2296)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction,
            5'- & 3'-end one pass sequencing: Departent of Virology and Human
            Genome Center, Institute of Medical Science, University of Tokyo
            (partly supported by Science and Technology Agency).
FEATURES             Location/Qualifiers
     source          1. .2296
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAT02564"
                     /cell_line="KATO III"
                     /cell_type="signet-ring cell carcinoma"
                     /clone_lib="KAT"
                     /note="cloning vector: pME18SFL3"
ORIGIN
```

Query Match 32.0%; Score 671; DB 8; Length 2296;
Best Local Similarity 100.0%; Pred. No. 2.8e-106;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1360 GAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAA 1419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAA 60

Qy       1420 GAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATC 1479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATC 120

Qy       1480 AGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCCACCCGACTACTTGGAA 1539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCCACCCGACTACTTGGAA 180

Qy       1540 TGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACC 1599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACC 240

Qy       1600 TTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCA 1659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCA 300

Qy       1660 AAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTG 1719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTG 360

Qy       1720 ACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCC 1779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCC 420

Qy       1780 TCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGG 1839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGG 480

Qy       1840 GAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCT 1899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCT 540

Qy       1900 GTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATG 1959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATG 600

Qy       1960 ACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGT 2019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGT 660

Qy       2020 TTCAATCTTTA 2030
             ||||||||||||
Db        661 TTCAATCTTTA 671
```

```
RESULT 11
BC061676
LOCUS       BC061676                3616 bp    mRNA    linear   VRT 03-FEB-2004
DEFINITION  Xenopus laevis hypothetical protein MGC68822, mRNA (cDNA clone
            MGC:68822 IMAGE:4633564), complete cds.
ACCESSION   BC061676
VERSION     BC061676.1  GI:38197625
KEYWORDS    MGC.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 3616)
  AUTHORS   Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
            and Richardson,P.
  TITLE     Genetic and genomic tools for Xenopus research: The NIH Xenopus
            initiative
  JOURNAL   Dev. Dyn. 225 (4), 384-391 (2002)
  PUBMED    12454917
REFERENCE   2  (bases 1 to 3616)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   3  (bases 1 to 3616)
  AUTHORS   Klein,S. and Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2003) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
  REMARK    NIH-MGC Project
COMMENT     Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Igor Dawid
            cDNA Library Preparation: Life Technologies, Inc.
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 128 Row: p Column: 13.

FEATURES Location/Qualifiers

source 1. .3616
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:68822 IMAGE:4633564"
/tissue_type="Lung, adult Xenopus"
/clone_lib="NICHD_XGC_Lu1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene 1. .3616
/gene="MGC68822"
/db_xref="GeneID:398987"

CDS 115. .1641
/gene="MGC68822"
/codon_start=1
/product="MGC68822 protein"
/protein_id="AAH61676.1"
/db_xref="GI:38197626"
/db_xref="GeneID:398987"
/translation="MSRDADAVAKITESTYRNVMNQFNPGLRNLVNLGRNYEKAVAAM
VQAGRAYYDGVAKMGELASASHVSKELGHVLLEISIVHKRLNDNLEENFKKFHREIVA
ELERKTDQDIKYMNATLKRYQTEHKNKVDSLEKSQSDLKKLKRKSQGGRNVTKYEIKE
SEYVETITSRQSEINKFIAQGCREALLEEKRRFCFLVDKHCTFSNYINYFHVQAVDLL
TSKLPKWQETCSDISQVPDTVVNIMQDLKTPVSTPISASPEPSPLMDKKALMAGDTLY
RNTLRMPPAPQTKAQTSPLVDMFNKPPVVQKLTFDNPRTSQGESQEEGTLPRSMSVAT
GLNQMKKRKVKTIFPHSGNGSTQMSFSQGDIITLLIPEEKDGWLYGEHDTTKVKGWFP
SSYTRPLEENGREPLKVPSVSPSPAPVRSLSTANLVEKSGVVLPEPDYLEPMSGKRSS
SSNNSNRGYTFSPLIAKDPPPSTANGITNHPFLSGENPFSTIKLRPTVTNDRSAPIIR
"

misc_feature 1132. .1311
/gene="MGC68822"
/note="SH3; Region: Src homology 3 domains"
/db_xref="CDD:smart00326"

ORIGIN

Query Match 28.4%; Score 594.4; DB 5; Length 3616;
Best Local Similarity 65.1%; Pred. No. 5.7e-93;
Matches 1014; Conservative 0; Mismatches 511; Indels 33; Gaps 8;

```
Qy   208 GCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGG 267
Db   106 GACGTAGCCATGTCCCGGGACGCAGACGCCGTGGCAAAGATCACAGAGAGCACGTACAGG 165

Qy   268 AATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTAT 327
Db   166 AATGTTATGAATCAGTTCAACCCTGGGTTAAGAAACTTGGTTAATCTTGGAAGGAACTAT 225

Qy   328 GAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAG 387
Db   226 GAAAAAGCAGTGGCAGCAATGGTGCAGGCAGGAAGAGCGTATTATGATGGAGTAGCCAAG 285

Qy   388 ATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAG 447
Db   286 ATGGGCGAATTGGCCTCTGCATCTCATGTATCAAAAGAGCTAGGCCACGTACTTCTGGAG 345

Qy   448 ATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 507
Db   346 ATTTCCATCGTTCACAAAAGGCTGAATGACAACTTAGAAGAAAATTTTAAAAAATTCCAC 405

Qy   508 AAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCA 567
Db   406 AGAGAGATTGTTGCAGAGCTGGAGAGAAAGACTGATCAAGATATAAAGTATATGAATGCC 465

Qy   568 ACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAA 627
Db   466 ACCCTTAAGAGGTACCAGACAGAGCACAAAAATAAAGTAGACTCATTGGAGAAATCTCAA 525

Qy   628 GCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAA 687
Db   526 TCAGATCTAAAGAAACTCAAGCGAAAAAGTCAAGGAGGAAGAAACGTCACTAAATATGAA 585

Qy   688 CACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 747
Db   586 ATTAAAGAATCCGAGTATGTGGAAACCATAACGTCCAGACAGTCGGAGATAAATAAATTC 645

Qy   748 ATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTT 807
Db   646 ATAGCGCAGGGCTGCAGGGAGGCTCTGCTTGAGGAGAAGAGACGATTCTGTTTCTTAGTG 705

Qy   808 GATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTA 867
Db   706 GACAAACACTGCACCTTCTCGAATTATATCAACTATTTCCACGTACAGGCTGTGGATCTA 765

Qy   868 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAG 927
Db   766 CTCACCTCCAAGCTGCCCAAGTGGCAGGAGACGTGTAGTGACATCTCTCAGGTTCCAGAC 825

Qy   928 AAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACT 987
Db   826 ACAGTCGTTAATATTATGCAGGACCTGAAAACTCCAGTGTCCACACCAATCTCTGCTTCC 885

Qy   988 CCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTT 1047
Db   886 CCTGAACCATCCCCATTAATGGA------CAAAAAGGCCCTGATGGCTGGAGATACCCTG 939

Qy  1048 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTG 1107
Db   940 TATAGGAACACACTAAGGATGCCTCCAGCACCCCAGACGAAAGCGCAGACGAGTCCCCTG 999

Qy  1108 ATCGATATGTTTAATAACCCAGCCACGGCTG-----CCCCGAATTCACAAAGGGTAAATA 1162
Db  1000 GTTGACATGTTTAACAAACCACCAGTTGTCCAAAAACTAACATTTGACAATCCAAGGACA 1059

Qy  1163 ATTCAA-CAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGA 1221
Db  1060 TCACAAGGAGAGAGTCAAGAAGAAGGCACTTTACCTCGCTCCATGTCCGTAGCAACAGGT 1119

Qy  1222 CTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAAC 1281
Db  1120 TTGAATCAAATGAAAAAACGAAAAGTCAAAACCATCTTCCCACACA---GCGGAAACGGC 1176

Qy  1282 AAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAG 1341
Db  1177 AGCACACAGATGAGCTTTTCTCAGGGCGACATCATCACACTTCTTATCCCAGAGGAAAAG 1236

Qy  1342 GATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCG 1401
Db  1237 GATGGCTGGCTGTACGGAGAGCATGATACAACCAAAGTTAAAGGTTGGTTCCCATCATCC 1296

Qy  1402 TACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCA------CGCCA 1455
Db  1297 TATACAAGGCCGCTGGAAGAAAACGGCAGAGAGCCATTGAAAGTGCCAAGCGTGTCCCCA 1356

Qy  1456 AGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTC 1515
Db  1357 AGTCCTGCTCCAGTCCGGAGTCTCAGCACAGCCAACTTAGTTGAAAAAAGCGGAGTTGTC 1416
```

```
Qy       1516 ATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCA 1575
Db       1417 CTCCCAGAACCGGATTATTTGGAACCAATGTC--TGGGAAACGCAGTAGCAGCAGCAACA 1474

Qy       1576 GATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCT 1635
Db       1475 ATAGCAACCGGGGATACA-----CTTTCAGCCCAC-----TCATTGCCAAAGATCCTCCT 1524

Qy       1636 CCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCC 1695
Db       1525 CCTTCAACTGCAAATGGCATCACTAACCATCCATTTCTAAGTGGAGAAAACCCATTTTCC 1584

Qy       1696 ACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAG 1753
Db       1585 ACCATCAAGCTCCGACCAACCGTAACAAATGACCGATCTGCGCCAATCATCCGATGAG 1642
```

RESULT 12
AX379210
LOCUS AX379210 559 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 252 from Patent WO0196389.
ACCESSION AX379210
VERSION AX379210.1 GI:19575050
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meagher,M.J., King,G.E., Xu,J. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196389-A 252 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .559
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 26.5%; Score 556; DB 6; Length 559;
Best Local Similarity 99.5%; Pred. No. 2.3e-86;
Matches 556; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        863 AACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGC 922
Db          1 AACTACTGAATTCCAAGCTGCCTCGGNGGCAGGAGACCTGTGTTGATGCCATCAAAGTGC 60

Qy        923 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTG 982
Db         61 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTG 120

Qy        983 GAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACA 1042
Db        121 GAACTCCTCANNCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACA 180

Qy       1043 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTC 1102
Db        181 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTC 240

Qy       1103 CCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATA 1162
Db        241 CCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATA 300

Qy       1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 1222
Db        301 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 360

Qy       1223 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 1282
Db        361 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 420
```

```
Qy       1283 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 1342
Db        421 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 480

Qy       1343 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 1402
Db        481 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 540

Qy       1403 ACACGAAGTTGCTGGAAGA 1421
Db        541 ACACGAAGTTGCTGGAAGA 559
```

RESULT 13
BX931244
LOCUS BX931244 685 bp mRNA linear VRT 02-FEB-2004
DEFINITION Gallus gallus finished cDNA, clone ChEST261d10.
ACCESSION BX931244
VERSION BX931244.1 GI:41631772
KEYWORDS .
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 685)
AUTHORS Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E., Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V., Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R., Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G., Tickle,C. and Wilson,S.A.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA sequencing project.
This sequence is from the BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection, from a library constructed by Elizabeth Bosch. cDNA was prepared from RNA extracted from ovary, normalised, and poly A-trimmed. EcoRI-NotI cut cDNA was then ligated into the vector. Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia coli DH10B.
FEATURES Location/Qualifiers
source 1. .685
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="ChEST261d10"
/clone_lib="CSEQRBN21"
/dev_stage="adult"
ORIGIN

Query Match 18.9%; Score 397; DB 5; Length 685;
Best Local Similarity 73.7%; Pred. No. 9.2e-59;
Matches 505; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

```
Qy        174 CCAGCCGCGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGA 233
Db          1 CCCGAAGCGCTCTGCGGAGAACGAGGCTGTCCCCGCCGGCACCATGTCGCGGGACCCCGA 60

Qy        234 GGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGG 293
Db         61 GGAGGTGAACAAGCTAACGGAGAGCACTTACAAAAATGTTATGGAGCAATTCAATCCAGG 120

Qy        294 GCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCT 353
Db        121 ACTGCGGAATTTAATAAATCTTGGGAAGAATTACGAAAAAGCTGTTAATGCTATGGTAGT 180

Qy        354 GGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCC 413
```

```
Db        181 GGCTGGAAGATCATACTACGACAGTCTGGCAAAGATTGGGGATATATCAGCAGATTCACC 240

Qy        414 CGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAA 473

Db        241 TGTTTCTAAAGAGTTAGGCCAAGTCCTGGTGGAAATTTCAAGAACGCACAAGAAACTTAA 300

Qy        474 CGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAA 533

Db        301 CGACAGCCTGGAGGAAAGTTTAAGAAATTTCATAAGGAAATTATATCTGAACTGGAGAA 360

Qy        534 GAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACA 593

Db        361 GAAAACAGACCTGGATGTAAAATACATGACCGCGACTTTAAAGAGGTACCAAACCGAACA 420

Qy        594 CAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAA 653

Db        421 CAGAAGTAAATTGGATTCCCTGGAGAAGTCTCAGGCTGAGCTGAAAAAAATCCGAAGGAA 480

Qy        654 AAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGAC 713

Db        481 AAGCCAAGGAGCACGAAATGTAATGAAATACGAGCATAAGGAAATGGAGTATTTGGAAAC 540

Qy        714 CGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCT 773

Db        541 CGTGAGCTCTCGACAGACTGACATTCAGAGATTTATTGCTGAAGGCTGTAGAGAAGCTCT 600

Qy        774 GCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCA 833

Db        601 ACTTGAAGAAAAAAGAAGATTCTGTTTTCTAGTTGACAAGCACTGCAGTTTTACCCAGCA 660

Qy        834 CATACATTATTATCACTTACAGTCT 858

Db        661 CATGCACTTCTATCACGCTCAGTGT 685
```

RESULT 14
BC068330
LOCUS BC068330 2563 bp mRNA linear VRT 30-MAR-2005
DEFINITION Danio rerio BAI1-associated protein 2-like 1, mRNA (cDNA clone MGC:85624 IMAGE:6525249), complete cds.
ACCESSION BC068330
VERSION BC068330.1 GI:46249678
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 2563)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRTM Mammalian Gene Collection Program Team
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2563)
AUTHORS .
CONSRTM NIH MGC Project
TITLE Direct Submission
JOURNAL Submitted (02-APR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

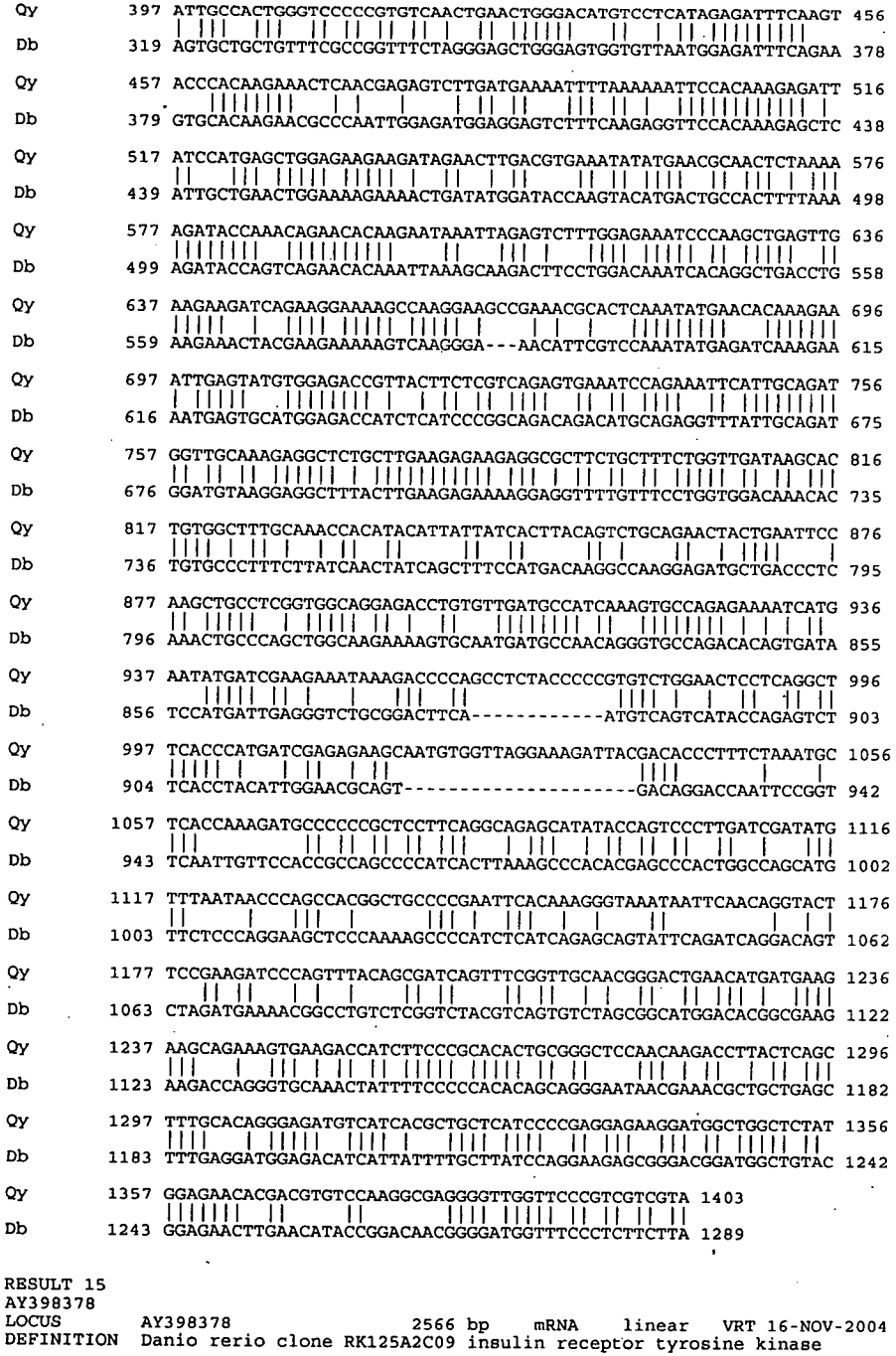
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 165 Row: b Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 38708008.

FEATURES Location/Qualifiers
source 1. .2563
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:85624 IMAGE:6525249"
/tissue_type="Kidney, zebrafish"
/clone_lib="NCI_CGAP_ZKid1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
gene 1. .2563
/gene="baiap2l1"
/note="synonyms: cb1023, zgc:85624, loc55971"
/db_xref="GeneID:387261"
/db_xref="ZFIN:ZDB-GENE-031201-1"
CDS 139. .1596
/gene="baiap2l1"
/codon_start=1
/product="baiap2l1 protein"
/protein_id="AAH68330.1"
/db_xref="GI:46249679"
/db_xref="GeneID:387261"
/db_xref="ZFIN:ZDB-GENE-031201-1"
/translation="MSRASEDVNKLTESTYKNVMEQFNPGLRNLVNLGKNYEKSVAAM
TFAGKAYFDALSKIGESAAVSPVSRELGVVLMEISEVHKNAQLEMEESFKRFHKELIA
ELEKKTDMDTKYMTATFKRYQSEHKLKQDFLDKSQADLKKLRRKSQGKHSSKYEIKEN
ECMETISSRQTDMQRFIADGCKEALLEEKRRFCFLVDKHCALSYQLSAFHDKAKEMLT
LKLPSWQEKCNDANRVPDTVISMIEGLRTSMSVIPESSPTLERSDRTNSGSIVPPPAP
SLKAHTSPLASMFSQEAPKSPISSEQYSDQDSLDENGLSRSTSVSSGMDTAKKTRVQT
IFPHTAGNNETLLSFEDGDIIILLIQEERDGWLYGELEHTGQRGWFPSSYCRSYNETT
IPNSRMESPFSSHIDLPEQQEEDEEVEEPVLLPPPDYSDSSRDRPIKPSTPSSQSTVS
LANGHSMPPFLGGGNPFATVRLRPTVTNDRSAPII"

ORIGIN

Query Match 18.3%; Score 382.6; DB 5; Length 2563;
Best Local Similarity 60.0%; Pred. No. 3.2e-56;
Matches 712; Conservative 0; Mismatches 439; Indels 36; Gaps 3;

```
Qy        217 ATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATG 276

Db        139 ATGTCTCGAGCCTCGGAGGACGTGAACAAACTCACCGAGAGCACATACAAGAATGTGATG 198

Qy        277 GAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCT 336

Db        199 GAACAGTTCAACCCAGGACTGCGGAATCTGGTTAACCTGGGGAAGAACTATGAGAAATCA 258

Qy        337 GTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAG 396

Db        259 GTTGCAGCAATGACCTTTGCCGGAAAGGCTTATTTTGATGCACTCTCAAAAATAGGGGAA 318
```



RESULT 15
AY398378
LOCUS AY398378 2566 bp mRNA linear VRT 16-NOV-2004
DEFINITION Danio rerio clone RK125A2C09 insulin receptor tyrosine kinase substrate (LOC55971) mRNA, complete cds.
ACCESSION AY398378
VERSION AY398378.1 GI:37681866
KEYWORDS .
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 2566)
AUTHORS Song,H.D., Sun,X.J., Deng,M., Zhang,G.W., Zhou,Y., Wu,X.Y., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Zon,L.I., Kanki,J.P., Liu,T.X., Look,A.T. and Chen,Z.
TITLE Hematopoietic gene expression profile in zebrafish kidney marrow
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (46), 16240-16245 (2004)
PUBMED 15520368
REFERENCE 2 (bases 1 to 2566)
AUTHORS Zhang,G.W., Sun,X.J., Wu,X.Y., Song,H.D., Zhou,Y., Liu,T.X., Deng,M., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Zon,L.I., Kanki,J.P., Look,A.T. and Chen,Z.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-2003) State Key Lab for Medical Genomics, Shanghai Institute of Hematology, Ruijin Hospital Affiliated to Shanghai Second Medical University, 197 Rui Jin Road II, Shanghai 200025, P. R. China
FEATURES Location/Qualifiers
source 1. .2566
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="RK125A2C09"
/tissue_type="kidney marrow"
gene 1. .2566
/gene="LOC55971"
CDS 156. .1613
/gene="LOC55971"
/codon_start=1
/product="insulin receptor tyrosine kinase substrate"
/protein_id="AAQ97811.1"
/db_xref="GI:37681867"
/translation="MSRASEDVNKLTESTYKNVMEQFNPGLRNLVNLGKNYEKSVAAM
TFAGKAYFDALSKIGESAAVSPVSRELGVVLMEISEVHKNAQLEMEESFKRFHKELIA
ELEKKTDMDTKYMTATFKRYQSEHKLKQDFLDKSQADLKKLRRKSQGKHSSKYEIKEN
ECMETISSRQTDMQRFIADGCKEALLEEKRRFCFLVDKHCALSYQLSAFHDKAKEMLT
VKLPSWQEKCNDATRVPDTVISMIEGLRTSMSVIPESSPTLERSDRTNSGSIVPPPAP
SLKAHTSPLASMFSQEAPKSPISSEQYSDQDSLDENGLSRSTSVSSGMDTAKKTRVQT
IFPHTAGNNETLLSFEDGDIIILLIQEERDGWLYGELEHTGQRGWFPSSYCRSYNETT
IPNSRMESPFSSHIDLPEQQEEDEEVEEPVLLPPPDYSDSSRDRPIKPSTPSSQSTVS
LANGHSMPPFLGGGNPFATVRLRPTVTNDRSAPII"
ORIGIN

Query Match 18.2%; Score 381; DB 5; Length 2566;
Best Local Similarity 59.9%; Pred. No. 6e-56;
Matches 711; Conservative 0; Mismatches 440; Indels 36; Gaps 3;

```
Db        396 GTGCACAAGAACGCCCAATTGGAGATGGAGGAGTCTTTCAAGAGGTTCCACAAAGAGCTC 455

Qy        517 ATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAA 576
              ||   ||| ||||| ||||| |  ||  | ||    || || ||||  || ||| | |||
Db        456 ATTGCTGAACTGGAAAAGAAAACTGATATGGATACCAAGTACATGACTGCCACTTTTAAA 515

Qy        577 AGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTG 636
              |||||||| |||||||||| ||  ||| |  |||| ||||| || ||||| ||
Db        516 AGATACCAGTCAGAACACAAATTAAAGCAAGACTTCCTGGACAAATCACAGGCTGACCTG 575

Qy        637 AAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAA 696
              |||·|  |  |||| ||||| ||||| |    | |  |   |||||||||   |||||||
Db        576 AAGAAACTACGAAGAAAAAGTCAAGGGA---AACATTCGTCCAAATATGAGATCAAAGAA 632

Qy        697 ATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGAT 756
              | ||||||  |||||||| |  | || || ||||  || || ||||  || |||||||||
Db        633 AATGAGTGCATGGAGACCATCTCATCCCGGCAGACAGACATGCAGAGGTTTATTGCAGAT 692

Qy        757 GGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCAC 816
              || || || |||||| | || |||||||| ||| | || || || ||||| || || |||
Db        693 GGATGTAAGGAGGCTTTACTAGAAGAGAAAAGGAGGTTTTGTTTCCTGGTGGACAAACAC 752

Qy        817 TGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 876
              |||| | || |  | ||  ||      ||  ||      || |    ||  | ||||    |
Db        753 TGTGCCCTTTCTTATCAACTATCAGCTTTCCATGACAAGGCCAAGGAGATGCTGACCGTC 812

Qy        877 AAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATG 936
              || |||||  | ||||| || |  ||   |||||||| ||  |||||||| | | | ||
Db        813 AAACTGCCCAGCTGGCAAGAAAAGTGCAATGATGCCACCAGGGTGCCAGACACAGTGATA 872

Qy        937 AATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCT 996
              |||||| || |   |  |||  ||              ||||·|  |  ||  || ||
Db        873 TCCATGATTGAGGGTCTGCGGACTTCA-----------ATGTCAGTCATACCAGAGTCT 920

Qy        997 TCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGC 1056
              ||||| |   | ||  | ||                              ||||       |   |
Db        921 TCACCTACATTGGAACGCAGT---------------------GACAGGACCAATTCCGGT 959

Qy       1057 TCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATG 1116
              |||          || || || || |||   | |||  | || || ||  ||  |   |||
Db        960 TCAATTGTTCCACCGCCAGCCCCATCACTTAAAGCCCACACGAGCCCACTGGCCAGCATG 1019

Qy       1117 TTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACT 1176
              ||     |  ||| |      ||| |  |||  |  |    ||          | | |
Db       1020 TTCTCCCAGGAAGCTCCCAAAAGCCCCATCTCATCAGAGCAGTATTCAGATCAGGACAGT 1079

Qy       1177 TCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAG 1236
                || ||   | |  |     || ||    || ||  | |  ||  || ||| |  |||
Db       1080 CTAGACGAAAACGGCCTGTCTCGGTCTACGTCAGTGTCTAGCGGCATGGACACGGCGAAA 1139

Qy       1237 AAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGC 1296
              |||   |  ||| | || || ||||| ||||| || ||    ||| | ||  | || |||
Db       1140 AAGACCAGGGTGCAAACTATTTTCCCCCACACAGCAGGGAATAACGAAACGCTGCTGAGC 1199

Qy       1297 TTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTAT 1356
              ||||  | |||||  |||| |   |||| |||| || |||  ||| || ||||| ||
Db       1200 TTTGAGGATGGAGACATCATTATTTTGCTTATCCAGGAAGAGCGGGACGGATGGCTGTAC 1259

Qy       1357 GGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTA 1403
              |||||||  ||     ||       |||| ||||| ||||| || ||
Db       1260 GGAGAACTTGAACATACCGGACAACGGGGATGGTTTCCGTCTTCTTA 1306
```

Search completed: March 20, 2006, 16:31:42
Job time : 10294 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:44:26 ; Search time 2596 Seconds
(without alignments)
1882.610 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.....aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	357.4	17.1	516	9	US-11-021-492-430	Sequence 430, App
	2	266.6	12.7	2078	12	US-11-128-061-1100	Sequence 1100, Ap
	3	266.6	12.7	2078	12	US-11-128-049-1100	Sequence 1100, Ap
	4	209.6	10.0	303	12	US-11-128-061-2130	Sequence 2130, Ap
	5	209.6	10.0	303	12	US-11-128-061-5772	Sequence 5772, Ap
	6	209.6	10.0	303	12	US-11-128-049-2130	Sequence 2130, Ap
	7	209.6	10.0	303	12	US-11-128-049-5772	Sequence 5772, Ap
c	8	147.2	7.0	640	6	US-09-925-065A-730687	Sequence 730687,
	9	81.4	3.9	1395	8	US-10-955-054A-111	Sequence 111, App
	10	79.4	3.8	1400	12	US-11-128-061-4742	Sequence 4742, Ap
	11	79.4	3.8	1400	12	US-11-128-049-4742	Sequence 4742, Ap
c	12	79.4	3.8	2036	8	US-10-996-217A-8	Sequence 8, Appli
	13	78.8	3.8	1400	12	US-11-136-527-7203	Sequence 7203, Ap
	14	78.8	3.8	1967	12	US-11-136-527-3107	Sequence 3107, Ap
	15	78	3.7	2120	8	US-10-689-742-163	Sequence 163, App
	16	77.8	3.7	932	9	US-11-031-206-65	Sequence 65, Appl
	17	77.4	3.7	1133	8	US-10-663-794-1	Sequence 1, Appli
	18	76	3.6	1400	12	US-11-136-527-6433	Sequence 6433, Ap
	19	76	3.6	3801	12	US-11-136-527-2337	Sequence 2337, Ap
	20	75.8	3.6	2255	9	US-11-096-568A-2260	Sequence 2260, Ap
	21	75.6	3.6	2797	12	US-11-120-308-133	Sequence 133, App
	22	75.6	3.6	4330	12	US-11-091-883-182	Sequence 182, App
c	23	75.4	3.6	1456	9	US-11-096-568A-12004	Sequence 12004, A
	24	75.4	3.6	2091	8	US-10-276-233A-19	Sequence 19, Appl
	25	75	3.6	1273	9	US-11-096-568A-12247	Sequence 12247, A
	26	75	3.6	4339	8	US-10-909-125-801	Sequence 801, App
c	27	74.8	3.6	2162	12	US-11-091-883-330	Sequence 330, App
	28	74.6	3.6	600	12	US-11-123-896-256	Sequence 256, App
	29	74.6	3.6	755	7	US-10-973-115B-153	Sequence 153, App
	30	74.6	3.6	755	8	US-10-131-826A-153	Sequence 153, App
	31	74.6	3.6	3082	9	US-11-096-568A-2443	Sequence 2443, Ap
	32	74.4	3.5	600	12	US-11-136-527-7965	Sequence 7965, Ap
	33	74.4	3.5	625	9	US-11-096-568A-2599	Sequence 2599, Ap
c	34	74.4	3.5	804	12	US-11-091-883-265	Sequence 265, App
	35	74.4	3.5	2183	12	US-11-000-688-485	Sequence 485, App
	36	74.4	3.5	2924	12	US-11-136-527-3869	Sequence 3869, Ap
	37	74.4	3.5	3501	7	US-10-973-115B-37	Sequence 37, Appl
	38	74.4	3.5	3501	8	US-10-131-826A-37	Sequence 37, Appl
c	39	74	3.5	850	9	US-11-096-568A-3385	Sequence 3385, Ap
	40	74	3.5	1358	9	US-11-096-568A-2642	Sequence 2642, Ap
	41	74	3.5	2537	12	US-11-102-978-8	Sequence 8, Appli
	42	73.8	3.5	799	12	US-11-195-459-3	Sequence 3, Appli
c	43	73.8	3.5	807	9	US-11-096-568A-706	Sequence 706, App
	44	73.8	3.5	1660	9	US-11-096-568A-2748	Sequence 2748, Ap
	45	73.8	3.5	2476	7	US-10-973-115B-489	Sequence 489, App

ALIGNMENTS

RESULT 1
US-11-021-492-430
; Sequence 430, Application US/11021492
; Publication No. US20060031947A1
; GENERAL INFORMATION:
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0368-USA
; CURRENT APPLICATION NUMBER: US/11/021,492
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/307,670
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 698
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 40, 156, 175, 190, 216, 291, 344, 370, 431
; OTHER INFORMATION: n = A,T,C or G
US-11-021-492-430

Query Match 17.1%; Score 357.4; DB 9; Length 516;
Best Local Similarity 83.9%; Pred. No. 7.2e-35;
Matches 433; Conservative 0; Mismatches 80; Indels 3; Gaps 3;

```
Qy        429 GGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGA 488
              | | | |||||||||||||||||| || |||| | ||| ||||| |||| |||||||||
Db          1 GAGGCCTGTCCTCATAGAGATTTCCAGCACCCCAAGGAANCTCAATGAGACTCTTGATGA 60

Qy        489 AAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGA 548
              |||||| |||||||||| || || |||||||||||||||||||||||.|||||||||
Db         61 AAATTTCAAAAAATTCCATAAGGATATTATCCATGAGCTGGAGAAGAAGACAGAACTTGA 120

Qy        549 CGTGAAATATATGAA-CGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATT-A 606
              || || |||||||| ||| ||||| |||||||| || |||| |||| ||||| ||| |
Db        121 TGTAAAGTATATGAACCGCTACTCTTAAAAGATATNAAGCAGAGCACAGGAATANATTAA 180
```

```
Qy        607 GAGTCTTTGGAGAAATCCCAAGC-TGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAG 665
              || |||||| ||||||| || ||  ||| |||||| ||||||||| ||||| |||||  |
Db        181 GATTCTTTGNAGAAATCTCAGGCAGGAGCTGAAGANGATCAGAAGAAAAAGTCAAGGTGG 240

Qy        666 CCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCG 725
              |||||||||||||||| || |||||||||||||||||||||||||  ||  |||||
Db        241 ACGAAACGCACTCAAATACGAGCACAAAGAAATTGAGTATGTGGAGACGGNGACGTCTCG 300

Qy        726 TCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAA 785
              |||||||| |||||||| ||||| |||||||  | |||||||| || ||
Db        301 CCAGAGTGAGATCCAGAAGTTCATTGCCGATGGCTGCAAAGAANCCCTGCTTGAGGAAAA 360

Qy        786 GAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTA 845
              |||||||| |||||||||| || |||||||  |||||| | |||||||| || ||
Db        361 AAGGCGCTTNTGCTTTCTGGTGGACAAGCACTGCAGCTTTGCCAGTCACATACACTACTA 420

Qy        846 TCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGT 905
              |||| | ||| |||||||  |||| ||||||||||||| ||||||||||| ||||||
Db        421 TCACATGCAGNCTGCAGAGTTACTTAATTCCAAGCTGCCCCGGTGGCAGGAAACCTGTTG 480

Qy        906 TGATGCCATCAAAGTGCCAGAGAAAATCATGAATAT 941
              ||||||||  ||||| ||||||||||||||||| ||
Db        481 TGATGCCACAAAAGTCCCAGAGAAAATCATGAACAT 516


RESULT 2
US-11-128-061-1100
; Sequence 1100, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
;  TITLE OF INVENTION:  TO MONITOR GENE EXPRESSION
;  FILE REFERENCE: 01997.027701
;  CURRENT APPLICATION NUMBER: US/11/128,061
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1100
;   LENGTH: 2078
;   TYPE: DNA
;   ORGANISM: Cricetinae
US-11-128-061-1100

  Query Match             12.7%;  Score 266.6;  DB 12;  Length 2078;
  Best Local Similarity   59.2%;  Pred. No. 5.3e-24;
  Matches  455;  Conservative    0;  Mismatches  314;  Indels    0;  Gaps    0;

Qy        230 CCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATC 289
              | |||||| || |||||||||||||  |    |||   ||   | ||||| ||||||||  |
Db         77 CGGAGGAGATGCACCGGCTCACGGAAAATGTCTATAAGACCATCATGGAGCAGTTCAACC 136

Qy        290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGA 349
              |  | || || || || ||   | ||||||| || || ||||||||  |    | | | |
Db        137 CCAGCCTCCGCAACTTCATCGCCATGGGGAAGAACTACGAGAAAGCCCTGGCAGGTGTCA 196

Qy        350 TCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGT 409
              | | || |  |||| ||| | |||||   |||  ||||| || ||| | ||||  | |
Db        197 CCTTCGCTGCCAAAGGCTATTTCGATGCTCTGGTAAAGATGGGGGAGCTGGCCAGCGAGA 256

Qy        410 CCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC 469
              ||  |  || |  ||||| ||  | ||||||  |  ||||  |      ||||  | |
Db        257 GCCAGGGCTCTAAGGAACTTGGGGACGTCCTCTTCCAGATGGCCGAAGTCCACCGGCAGA 316

Qy        470 TCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGG 529
              || |  |   ||  || || |  | ||    || ||||| ||| |  ||    ||||||
Db        317 TCCAGAACCAACTGGAAGAGATGCTAAAGTCTTTTCACAATGAGCTGCTCACACAGCTGG 376

Qy        530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAG 589
              |  ||||| |||| || |||   |  ||| |||  ||  | || || | ||||||  |||
Db        377 AACAGAAGGTAGAGCTGGACTCCAGGTATCTGAGTGCTGCACTGAAGAAATACCAGGCAG 436

Qy        590 AACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
              |  || | ||  |||  |||  || |||| || |  ||  |||||| |||||||| || | |
Db        437 AGCAGAGGAGCAAAGGAGATGCTCTGGACAAGTGTCAGGCTGAGCTGAAGAAGCTCCGTA 496

Qy        650 GGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGG 709
              |||  |||||||| |||   ||| | |  || ||     |||| ||  |  ||||  | |
Db        497 AGAAGAGCCAAGGGAGCAAGAACCCTCAGAAGTACTCGGACAAGGAGCTGCAGTACATTG 556

Qy        710 AGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGG 769
              |  || | |          ||| ||||  |  |||| | | |  | |||||  | |||  |
Db        557 ATGCCATCAGCAACAAGCAGGGTGAGCTGGAGAACTACGTATCTGATGGCTACAAGACAG 616

Qy        770 CTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAA 829
              | ||    || |||  |||| | |||||||| ||||| || ||||| ||||   | || |
Db        617 CACTCACCGAGGAGCGGAGGAGGTTCTGCTTCCTGGTGGAAAAGCAGTGTGCTGTGGCCA 676

Qy        830 ACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGT 889
              |  ||     |  ||| ||||   ||       || || ||       ||||||||   ||
Db        677 AGAACTCTGCTGCTTACCACTCCAAGGGCAAGGAGCTGCTAGCCCAGAAGCTGCCCGTGT 736

Qy        890 GGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAG 949
              |||||  |  |||||| |||  ||| |||  |  |||||          |  |  ||||  |
Db        737 GGCAGCAAGCCTGTGCTGACCCCAACAAGATTCCAGACCGCGCTGTCCAGCTGATGCAGC 796

Qy        950 AAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTC 998
              |  |||   | |     |   ||  | |  || | || |.    || ||
Db        797 AGATAGCCAGCAGCAACGGCTCCATCCTTCCTAGTACCCTGTCAGCCTC 845


RESULT 3
US-11-128-049-1100
; Sequence 1100, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FO
;  TITLE OF INVENTION:  MAKING AND USING SAME
;  FILE REFERENCE: 01997.027700
;  CURRENT APPLICATION NUMBER: US/11/128,049
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1100
;   LENGTH: 2078
;   TYPE: DNA
;   ORGANISM: Cricetinae
US-11-128-049-1100

  Query Match             12.7%;  Score 266.6;  DB 12;  Length 2078;
  Best Local Similarity   59.2%;  Pred. No. 5.3e-24;
  Matches  455;  Conservative    0;  Mismatches  314;  Indels    0;  Gaps    0;

Qy        230 CCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATC 289
              | |||||| || |||||||||||||  |    |||   ||   | ||||| ||||||||  |
Db         77 CGGAGGAGATGCACCGGCTCACGGAAAATGTCTATAAGACCATCATGGAGCAGTTCAACC 136
```

```
Qy        290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGA 349
              |  | || || || || ||   | ||||||| || || ||||||||  |     | | | |
Db        137 CCAGCCTCCGCAACTTCATCGCCATGGGGAAGAACTACGAGAAAGCCCTGGCAGGTGTCA 196

Qy        350 TCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGT 409
              | | || |  |||| ||| | |||||   |||  ||||| || ||| | ||||  | |
Db        197 CCTTCGCTGCCAAAGGCTATTTCGATGCTCTGGTAAAGATGGGGGAGCTGGCCAGCGAGA 256

Qy        410 CCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC 469
              ||  |  || |  ||||| ||  | |||||| |  |||| |       |||| | |
Db        257 GCCAGGGCTCTAAGGAACTTGGGGACGTCCTCTTCCAGATGGCCGAAGTCCACCGGCAGA 316

Qy        470 TCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGG 529
              ||  |   |    || || || |   | ||     || ||||| ||| |  ||    ||||||
Db        317 TCCAGAACCAACTGGAAGAGATGCTAAAGTCTTTTCACAATGAGCTGCTCACACAGCTGG 376

Qy        530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAG 589
              |  ||||| |||| || |||   |  ||| |||  ||  | || || | ||||||  |||
Db        377 AACAGAAGGTAGAGCTGGACTCCAGGTATCTGAGTGCTGCACTGAAGAAATACCAGGCAG 436

Qy        590 AACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
              | || | ||  |||  ||| || |||| || |  || |||||| |||||||| || | |
Db        437 AGCAGAGGAGCAAAGGAGATGCTCTGGACAAGTGTCAGGCTGAGCTGAAGAAGCTCCGTA 496

Qy        650 GGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGG 709
              |||  ||||||||  |||   ||| | |  || ||      |||| ||  |  ||||  | |
Db        497 AGAAGAGCCAAGGGAGCAAGAACCCTCAGAAGTACTCGGACAAGGAGCTGCAGTACATTG 556

Qy        710 AGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGG 769
              |  || | |      ||| |||| |  |||| | | |  | ||||| | |||    |
Db        557 ATGCCATCAGCAACAAGCAGGGTGAGCTGGAGAACTACGTATCTGATGGCTACAAGACAG 616

Qy        770 CTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAA 829
              | ||    || |||  |||| | |||||||| ||||| || ||||| ||||    | || |
Db        617 CACTCACCGAGGAGCGGAGGAGGTTCTGCTTCCTGGTGGAAAAGCAGTGTGCTGTGGCCA 676

Qy        830 ACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGT 889
              |  ||     |  ||| ||||   ||       || || ||        ||||||||   ||
Db        677 AGAACTCTGCTGCTTACCACTCCAAGGGCAAGGAGCTGCTAGCCCAGAAGCTGCCCGTGT 736

Qy        890 GGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAG 949
              |||||  |  |||||| |||  ||| |||  | |||||         |  |  ||||  |
Db        737 GGCAGCAAGCCTGTGCTGACCCCAACAAGATTCCAGACCGCGCTGTCCAGCTGATGCAGC 796

Qy        950 AAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTC 998
              | |||   | |      |    ||  | |  || | || |     || ||
Db        797 AGATAGCCAGCAGCAACGGCTCCATCCTTCCTAGTACCCTGTCAGCCTC 845
```

RESULT 4
US-11-128-061-2130
; Sequence 2130, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2130
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-2130

Query Match 10.0%; Score 209.6; DB 12; Length 303;
Best Local Similarity 83.1%; Pred. No. 4.3e-17;
Matches 250; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

```
Qy        674 CACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 733
              |||| |||||||| |||||||||||||||||||||||||| |||||||||| |||||||
Db          1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTCGCCAGAGTG 60

Qy        734 AAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCT 793
              ||||||| || |||||||||||||| |||||||| || |||||||| || || | |||||
Db         61 AAATCCATAAGTTCATTGCAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAAGCGCT 119

Qy        794 TCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTAC 853
              |||||||||||||| || |||||||| ||||||| || |||||||| || |||||| | |
Db        120 TCTGCTTTCTGGTGGACAAGCACTGTAGCTTTGCCAATCACATACACTACTATCACCTTC 179

Qy        854 AGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCA 913
              |||||||  || |||| |||||||| ||| | ||| ||||||| ||||||    |||||||
Db        180 AGTCTGCNAAATTACTTAATTCCAACCTGTCCCGGAGGCAGGAAACCTGTCGAGATGCCA 239

Qy        914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
                 |||  |||||||||||||||| ||||| |||| |||||  || ||   ||  || |
Db        240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy        974 C 974
              |
Db        300 C 300
```

RESULT 5
US-11-128-061-5772
; Sequence 5772, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5772
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-5772

Query Match 10.0%; Score 209.6; DB 12; Length 303;
Best Local Similarity 83.1%; Pred. No. 4.3e-17;
Matches 250; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

```
Qy        674 CACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 733
              |||| |||||||| |||||||||||||||||||||||||| |||||||||| |||||||
Db          1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTCGCCAGAGTG 60

Qy        734 AAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCT 793
              ||||||| || |||||||||||||| |||||||| || |||||||| || || | |||||
Db         61 AAATCCATAAGTTCATTGCAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAAGCGCT 119

Qy        794 TCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTAC 853
              |||||||||||| || ||||||||| ||||||| || |||||||| || |||||| | |
Db        120 TCTGCTTTCTGGTGGACAAGCACTGTAGCTTTGCCAATCACATACACTACTATCACCTTC 179

Qy        854 AGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCA 913
              |||||||  || |||| |||||||| ||| | ||| ||||||| ||||||   |||||||
Db        180 AGTCTGCNAAATTACTTAATTCCAACCTGTCCCGGAGGCAGGAAACCTGTCGAGATGCCA 239

Qy        914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
                  |||  |||||||||||||||| ||||| |||| |||||| || ||    ||  || |
Db        240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy        974 C 974
              |
Db        300 C 300
```

RESULT 6
US-11-128-049-2130
; Sequence 2130, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2130
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-2130

Query Match 10.0%; Score 209.6; DB 12; Length 303;
Best Local Similarity 83.1%; Pred. No. 4.3e-17;
Matches 250; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

```
Qy        674 CACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 733
              |||| |||||||| |||||||||||||||||||||||||| |||||||||| |||||||
Db          1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTCGCCAGAGTG 60

Qy        734 AAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCT 793
              ||||||| || |||||||||||||| |||||||| || |||||||| || || | |||||
Db         61 AAATCCATAAGTTCATTGCAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAAGCGCT 119

Qy        794 TCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTAC 853
              |||||||||||| || ||||||||| ||||||| || |||||||| || |||||| | |
Db        120 TCTGCTTTCTGGTGGACAAGCACTGTAGCTTTGCCAATCACATACACTACTATCACCTTC 179

Qy        854 AGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCA 913
              |||||||  || |||| |||||||| ||| | ||| ||||||| ||||||   |||||||
Db        180 AGTCTGCNAAATTACTTAATTCCAACCTGTCCCGGAGGCAGGAAACCTGTCGAGATGCCA 239

Qy        914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
                  |||  |||||||||||||||| ||||| |||| |||||| || ||    ||  || |
Db        240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy        974 C 974
              |
Db        300 C 300
```

RESULT 7
US-11-128-049-5772
; Sequence 5772, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FO
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5772
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-5772

Query Match 10.0%; Score 209.6; DB 12; Length 303;
Best Local Similarity 83.1%; Pred. No. 4.3e-17;
Matches 250; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

```
Qy        674 CACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 733
              |||| |||||||| |||||||||||||||||||||||||| |||||||||| |||||||
Db          1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTCGCCAGAGTG 60

Qy        734 AAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCT 793
              ||||||| || |||||||||||||| |||||||| || |||||||| || || | |||||
Db         61 AAATCCATAAGTTCATTGCAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAAGCGCT 119

Qy        794 TCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTAC 853
              |||||||||||| || ||||||||| ||||||| || |||||||| || |||||| | |
Db        120 TCTGCTTTCTGGTGGACAAGCACTGTAGCTTTGCCAATCACATACACTACTATCACCTTC 179

Qy        854 AGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCA 913
              |||||||  || |||| |||||||| ||| | ||| ||||||| ||||||   |||||||
Db        180 AGTCTGCNAAATTACTTAATTCCAACCTGTCCCGGAGGCAGGAAACCTGTCGAGATGCCA 239

Qy        914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
                  |||  |||||||||||||||| ||||| |||| |||||| || ||    ||  || |
Db        240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy        974 C 974
              |
Db        300 C 300
```

```
RESULT 8
US-09-925-065A-730687/c
; Sequence 730687, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 730687
;   LENGTH: 640
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-730687

  Query Match             7.0%;  Score 147.2;  DB 6;  Length 640;
  Best Local Similarity   92.3%;  Pred. No. 1.3e-09;
  Matches  155;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        709 GAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 GAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAG 581

Qy        769 GCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCA 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCA 521

Qy        829 AACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 876
              |||||||||||||||||||||||||||   |  |  |    |||| |
Db        520 AACCACATACATTATTATCACTTACAGGTGGGTGTGGTGGCAAATTGC 473


RESULT 9
US-10-955-054A-111
; Sequence 111, Application US/10955054A
; Publication No. US20050266420A1
; GENERAL INFORMATION:
;  APPLICANT: PUSZTAI, LAJOS
;  APPLICANT:  SYMMANS, W. FRASER
;  APPLICANT:  HESS, KENNETH R.
;  APPLICANT:  AYERS, MARK
;  APPLICANT:  STEC, JAMES
;  TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
;  FILE REFERENCE: UTXC:880US
;  CURRENT APPLICATION NUMBER: US/10/955,054A
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 195
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
;   LENGTH: 1395
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-955-054A-111

  Query Match             3.9%;  Score 81.4;  DB 8;  Length 1395;
  Best Local Similarity   70.3%;  Pred. No. 0.096;
  Matches  109;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1942 ATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCA 2001
              | || ||  ||  ||||      |        || ||   | | ||    ||| || | |
Db       1216 AGGAATAGCAGAATAATTGGTGCTTGTAAACATTTAAGATTCTCCTGTGGATTTTGGTGA 1275

Qy       2002 AGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2061
              | || ||||||   || |||  |  ||||||||||||||||||||||||||||||||||||
Db       1276 GTGATCATTAAACTGTTTTCCAACTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1335

Qy       2062 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||
Db       1336 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1370


RESULT 10
US-11-128-061-4742
; Sequence 4742, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
;  TITLE OF INVENTION:  TO MONITOR GENE EXPRESSION
;  FILE REFERENCE: 01997.027701
;  CURRENT APPLICATION NUMBER: US/11/128,061
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4742
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Cricetinae
US-11-128-061-4742

  Query Match             3.8%;  Score 79.4;  DB 12;  Length 1400;
  Best Local Similarity   54.5%;  Pred. No. 0.17;
  Matches  186;  Conservative    0;  Mismatches  146;  Indels    9;  Gaps    1;

Qy       1193 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA 1252
              | |  || |||   ||  | ||| | || ||  |     |           |  || ||
Db        467 TGCCACGTTCAAGCTCCATGGCAGCTGGCCTAGAACGTAACGGCCGCATGCGGGTCAAAG 526

Qy       1253 CCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATG 1312
              |||| ||| | ||| | || ||   |||||  ||  | ||| |||||     |||| ||
Db        527 CCATTTTCTCCCACGCGGCTGGTGACAACAGCACTCTGCTGAGCTTCAAGGAGGGCGACC 586

Qy       1313 TCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGT 1372
              ||||||| |||||  | || ||||     || |||||||  |||||| ||    ||
Db        587 TCATCACCCTGCTAGTGCCTGAGGCCCGTGACGGCTGGCACTATGGCGAGAGTGAGAAAA 646

Qy       1373 CCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAG 1432
              ||||   | |||| ||||| || |  || |||||   | | |||||     | |  | |
Db        647 CCAAAATGCGGGGCTGGTTTCCTTTCTCCTACACTCGGGTCCTGGACAGTGACGGAAGTG 706

Qy       1433 AAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACT 1492
              |    ||  | ||          ||||     |  |  | ||||  ||||||  |  |||
Db        707 ACAGGTTGCACATG---------AGCCTGCAGCAGGGCAAGAGCAGCAGCACGGGCAACC 757

Qy       1493 TGTCTGAGAATAGCAGTGTTGTCATCCCCCCCACCCGACTAC 1533
              |    || ||   |    | | | |||||||  || ||||||
Db        758 TTCTGGACAAGGACGACCTGGCCGTCCCCCCTCCTGACTAC 798


RESULT 11
US-11-128-049-4742
; Sequence 4742, Application US/11128049
```

```
; Publication No. US20060010513A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
;  TITLE OF INVENTION:  MAKING AND USING SAME
;  FILE REFERENCE: 01997.027700
;  CURRENT APPLICATION NUMBER: US/11/128,049
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4742
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Cricetinae
US-11-128-049-4742

  Query Match             3.8%;  Score 79.4;  DB 12;  Length 1400;
  Best Local Similarity   54.5%;  Pred. No. 0.17;
  Matches  186;  Conservative    0;  Mismatches  146;  Indels    9;  Gaps    1;

Qy       1193 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA 1252
             | |  || |||   ||  | ||| | || |    |          |   || ||
Db        467 TGCCACGTTCAAGCTCCATGGCAGCTGGCCTAGAACGTAACGGCCGCATGCGGGTCAAAG 526

Qy       1253 CCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATG 1312
             |||| ||| | ||| | || ||   |||||  ||  | || |||||     |||| ||
Db        527 CCATTTTCTCCCACGCGGCTGGTGACAACAGCACTCTGCTGAGCTTCAAGGAGGGCGACC 586

Qy       1313 TCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGT 1372
             ||||||| |||||  | || ||||     || |||||||| |||||| ||    ||
Db        587 TCATCACCCTGCTAGTGCCTGAGGCCCGTGACGGCTGGCACTATGGCGAGAGTGAGAAAA 646

Qy       1373 CCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAG 1432
             ||||   | |||| ||||| || |  || |||||   | | |||||     | |  | |
Db        647 CCAAAATGCGGGGCTGGTTTCCTTTCTCCTACACTCGGGTCCTGGACAGTGACGGAAGTG 706

Qy       1433 AAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACT 1492
             |     ||  | ||       ||||     |  |  |  |||| |||||| |  |||
Db        707 ACAGGTTGCACATG---------AGCCTGCAGCAGGGCAAGAGCAGCAGCACGGGCAACC 757

Qy       1493 TGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTAC 1533
             |    || ||   |    | | | ||||||| || ||||||
Db        758 TTCTGGACAAGGACGACCTGGCCGTCCCCCCTCCTGACTAC 798


RESULT 12
US-10-996-217A-8/c
; Sequence 8, Application US/10996217A
; Publication No. US20050266561A1
; GENERAL INFORMATION:
;  APPLICANT: Revivicor, Inc.
;  APPLICANT:  Wells, Kevin
;  TITLE OF INVENTION: Use of Interfering RNA in the Production of Transgenic Animals
;  FILE REFERENCE: 10785.105070 REV 1015 US
;  CURRENT APPLICATION NUMBER: US/10/996,217A
;  CURRENT FILING DATE:  2004-11-22
;  PRIOR APPLICATION NUMBER: 60/523,938
;  PRIOR FILING DATE: 2003-11-21
;  NUMBER OF SEQ ID NOS: 272
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
;   LENGTH: 2036
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Construct
US-10-996-217A-8

  Query Match             3.8%;  Score 79.4;  DB 8;  Length 2036;
  Best Local Similarity   73.7%;  Pred. No. 0.16;
  Matches  101;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       1960 ACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGT 2019
             ||||  |   |  ||||  |||||||||| | | |   ||  | | | |  ||
Db        668 ACCATTTACCCATTTTAGGATATTTTCTTGGTTTCCCTTCTCAGTGAAACTTATTTTTAG 609

Qy       2020 TTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2079
               |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 GTAATCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 549

Qy       2080 AAAAAAAAAAAAAAAA 2096
              ||||||||||||||||
Db        548 AAAAAAAAAAAAAAAA 532


RESULT 13
US-11-136-527-7203
; Sequence 7203, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7203
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-7203

  Query Match             3.8%;  Score 78.8;  DB 12;  Length 1400;
  Best Local Similarity   65.1%;  Pred. No. 0.2;
  Matches   82;  Conservative   20;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       1971 CATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
             | ||| |:| || : ||| || :| | :|||| || |||          :: ::::   :
Db        1254 CTTTTGAWACATGKACTTTCTGWTTTWYTCAARTAAAAGTTGGAAACTAYWRMWHVMMMM 1313

Qy       2031 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2090
             ::::::|:||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        1314 MMWMWMARAAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1373

Qy       2091 AAAAAA 2096
              ||||||
Db        1374 AAAAAA 1379


RESULT 14
US-11-136-527-3107
; Sequence 3107, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
```

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3107
;   LENGTH: 1967
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-3107

  Query Match             3.8%;  Score 78.8;  DB 12;  Length 1967;
  Best Local Similarity   65.1%;  Pred. No. 0.19;
  Matches   82;  Conservative   20;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       1971 CATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
             | ||| |:| || : ||| || :| | :||||: || |||         :: ::::   :
Db       1821 CTTTTGAWACATGKACTTTCTGWTTTWYTCAARTAAAAGTTGGAAACTAYWRMWHVMMMM 1880

Qy       2031 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2090
              ::::::|:||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1881 MMWMWMARAAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1940

Qy       2091 AAAAAA 2096
              ||||||
Db       1941 AAAAAA 1946


RESULT 15
US-10-689-742-163
; Sequence 163, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT:  McCoy, John  M
;  APPLICANT:  LaVallie, Edward R
;  APPLICANT:  Racie, Lisa A
;  APPLICANT:  Evans, Cheryl
;  APPLICANT:  Merberg, David
;  APPLICANT:  Treacy, Maurice
;  APPLICANT:  Spaulding, Vikki
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: 00766.000091.10
;  CURRENT APPLICATION NUMBER: US/10/689,742
;  CURRENT FILING DATE:  2003-10-22
;  PRIOR APPLICATION NUMBER: 09/746,783
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 231
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 163
;   LENGTH: 2120
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-689-742-163

  Query Match             3.7%;  Score 78;  DB 8;  Length 2120;
  Best Local Similarity   59.5%;  Pred. No. 0.24;
  Matches  132;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy       1875 CTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTT 1934
             ||   |||  | |  |     |  || | | | | | |     ||  || ||| | ||| |
Db       1887 CTATGTAAATATAAGACTACTGGCTTTGTGAGGGAATGTTTGTGCAAAATTTTTTCCTCT 1946

Qy       1935 AAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATT 1994
             ||  || |  |||  ||  | || |  |     |     | ||   |    ||  |
Db       1947 AATGTATAATAGTGTTAAATTGATTAAAAATCTTCCAGAATTAATATTCCCTTTTGTCAC 2006

Qy       1995 CTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAA 2054
              | || || | | | ||||     | |   | |     ||||||||||||||||||||
Db       2007 TTTTTGAAAACATAATAAATCATTTGTATCTGTGCCTTAAAAAAAAAAAAAAAAAAAAAA 2066

Qy       2055 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              |||||||||||||||||||||||||||||||||||||||||||
Db       2067 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2108


Search completed: March 20, 2006, 14:27:56

Job time : 2598 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:39:55 ; Search time 381 Seconds
(without alignments)
9778.922 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.....aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	259.6	12.4	2080	2	US-08-878-563A-2	Sequence 2, Appli
2	259.6	12.4	2080	3	US-09-270-117-2	Sequence 2, Appli
3	259.6	12.4	2080	3	US-09-046-572-2	Sequence 2, Appli
4	241.6	11.5	2328	3	US-09-046-572-4	Sequence 4, Appli
5	82.8	4.0	997	3	US-09-907-794A-376	Sequence 376, App
6	82.8	4.0	997	3	US-09-905-125A-376	Sequence 376, App
7	82.8	4.0	997	3	US-09-902-775A-376	Sequence 376, App
8	82.8	4.0	997	3	US-09-906-700-376	Sequence 376, App
9	82.8	4.0	997	3	US-09-903-603A-376	Sequence 376, App
10	82.8	4.0	997	3	US-09-904-920A-376	Sequence 376, App
11	82.8	4.0	997	3	US-09-909-064-376	Sequence 376, App
12	82.8	4.0	997	3	US-09-905-381A-376	Sequence 376, App
13	82.8	4.0	997	3	US-09-906-618-376	Sequence 376, App
14	82.8	4.0	997	3	US-09-906-646-376	Sequence 376, App
15	82.8	4.0	997	3	US-09-904-462-376	Sequence 376, App
16	82.8	4.0	997	3	US-09-902-736A-376	Sequence 376, App
17	82.8	4.0	997	3	US-09-906-722A-376	Sequence 376, App
18	81.2	3.9	1570	3	US-10-012-231A-291	Sequence 291, App
19	81.2	3.9	1570	3	US-10-015-389A-291	Sequence 291, App
20	81.2	3.9	1570	3	US-10-006-768A-291	Sequence 291, App
21	81.2	3.9	1570	3	US-10-015-671A-291	Sequence 291, App
22	81.2	3.9	1570	3	US-10-015-393A-291	Sequence 291, App
23	81.2	3.9	1570	3	US-10-011-833A-291	Sequence 291, App
24	81.2	3.9	1570	3	US-10-006-041A-291	Sequence 291, App
25	81.2	3.9	1570	3	US-10-012-064A-291	Sequence 291, App
26	80.8	3.9	687	3	US-09-907-907A-38	Sequence 38, Appl
27	80.2	3.8	1474	3	US-08-821-994-64	Sequence 64, Appl
28	79	3.8	2550	9	5258287-23	Patent No. 5258287
29	78.8	3.8	1172	2	US-07-945-288-9	Sequence 9, Appli
30	78.8	3.8	1172	2	US-08-462-831-9	Sequence 9, Appli
31	78.8	3.8	1172	2	US-08-461-809-9	Sequence 9, Appli
32	78.8	3.8	1172	2	US-08-461-441-9	Sequence 9, Appli
33	78.8	3.8	1172	6	PCT-US93-08518-9	Sequence 9, Appli
c 34	78.6	3.8	396	3	US-09-640-173-10	Sequence 10, Appl
c 35	78.6	3.8	396	3	US-09-640-173-53	Sequence 53, Appl
c 36	78.6	3.8	396	3	US-09-713-550-10	Sequence 10, Appl
c 37	78.6	3.8	396	3	US-09-713-550-53	Sequence 53, Appl
c 38	78.6	3.8	396	3	US-09-825-294-10	Sequence 10, Appl
c 39	78.6	3.8	396	3	US-09-825-294-53	Sequence 53, Appl
c 40	78.6	3.8	396	3	US-09-970-966-10	Sequence 10, Appl
c 41	78.6	3.8	396	3	US-09-970-966-53	Sequence 53, Appl
42	78.6	3.8	708	3	US-09-270-767-13081	Sequence 13081, A
43	78.2	3.7	985	3	US-09-832-129-19	Sequence 19, Appl
44	78.2	3.7	985	3	US-09-832-129-32	Sequence 32, Appl
45	78	3.7	2276	3	US-09-205-258-183	Sequence 183, App

ALIGNMENTS

```
RESULT 1
US-08-878-563A-2
; Sequence 2, Application US/08878563A
; Patent No. 5891674
; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.
;   APPLICANT: Lal, Preeti
;   APPLICANT: Shah, Purvi
;   TITLE OF INVENTION: INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Incyte Pharmaceuticals, Inc.
;     STREET: 3174 Porter Drive
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/878,563A
;     FILING DATE: Filed Herewith
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Billings, Lucy J.
;     REGISTRATION NUMBER: 36,749
;     REFERENCE/DOCKET NUMBER: PF-0323 US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-855-0555
;     TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2080 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     LIBRARY: BRSTNOT04
;     CLONE: 918158
US-08-878-563A-2

  Query Match 12.4%; Score 259.6; DB 2; Length 2080;
```

```
 Best Local Similarity   59.2%;  Pred. No. 1.6e-48;
 Matches  442;  Conservative    0;  Mismatches  304;  Indels    0;  Gaps    0;

Qy        230 CCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATC 289
              | |||||| || |||||||||||||| |    |||   ||   | ||||| ||||||||
Db        112 CAGAGGAGATGCACCGGCTCACGGAAAATGTCTATAAGACCATCATGGAGCAGTTCAACC 171

Qy        290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGA 349
              || | || || || || ||   | ||||||| ||||| ||||| ||  |   | | |||
Db        172 CTAGCCTCCGGAACTTCATCGCCATGGGGAAGAATTACGAGAAGGCACTGGCAGGTGTGA 231

Qy        350 TCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGT 409
                 ||||  |||| |||||  || |   |||  ||||| || ||| | ||||  | |
Db        232 CGTATGCAGCCAAAGGCTACTTTGACGCCCTGGTGAAGATGGGGGAGCTGGCCAGCGAGA 291

Qy        410 CCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC 469
              ||  |  || |  ||||| ||| | || ||| |  ||||  |       ||||| | |
Db        292 GCCAGGGCTCCAAAGAACTCGGAGACGTTCTCTTCCAGATGGCTGAAGTCCACAGGCAGA 351

Qy        470 TCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGG 529
              || |  |   || || ||||   | ||    || ||||| ||| |  |      ||||||
Db        352 TCCAGAATCAGCTGGAAGAAATGCTGAAGTCTTTTCACAACGAGCTGCTTACGCAGCTGG 411

Qy        530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAG 589
              || ||||| | || || |||   |  ||| |||  ||  | ||||| | |||||| || |
Db        412 AGCAGAAGGTGGAGCTGGACTCCAGGTATCTGAGTGCTGCGCTAAAGAAATACCAGACTG 471

Qy        590 AACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
              | || | ||  |||   ||  |  |||| || |  || |||||| |||||||  |  | |
Db        472 AGCAAAGGAGCAAAGGCGACGCCCTGGACAAGTGTCAGGCTGAGCTGAAGAAGCTTCGGA 531

Qy        650 GGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGG 709
               ||| ||||| || |||   ||  | |  || ||     |||| ||  |  |||| | |
Db        532 AGAAGAGCCAGGGCAGCAAGAATCCTCAGAAGTACTCGGACAAGGAGCTGCAGTACATCG 591

Qy        710 AGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGG 769
              |  || | |         ||| | ||  |  |||| | | |  | || || | |||    |
Db        592 ACGCCATCAGCAACAAGCAGGGCGAGCTGGAGAATTACGTGTCCGACGGCTACAAGACCG 651

Qy        770 CTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAA 829
              | |||   || |||  |||||||||||||| ||||| || ||||| || | | | || |
Db        652 CACTGACAGAGGAGTGCAGGCGCTTCTGCTTCCTGGTGGAGAAGCAGTGCGCCGTGGCCA 711

Qy        830 ACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGT 889
              | ||         || |||   ||      || || |||        |||||||| | ||
Db        712 AGAACTCCGCGGCCTACCACTCCAAGGGCAAGGAGCTGCTGCCGCAGAAGCTGCCGCTGT 771

Qy        890 GGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAG 949
              ||||  || ||||||  ||  ||| |||  | || |||       || |  | ||  |
Db        772 GGCAACAGGCCTGTGCCGACCCCAGCAAGATCCCGGAGCGCGCGGTGCAGCTCATGCAGC 831

Qy        950 AAATAAAGACCCCAGCCTCTACCCCC 975
              |  |    | |   | | | |||| |
Db        832 AGGTGGCCAGCAACGGCGCCACCCTC 857


RESULT 2
US-09-270-117-2
; Sequence 2, Application US/09270117
; Patent No. 6265550
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/270,117
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/878,563
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0323 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2080 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT04
;      CLONE:  918158
US-09-270-117-2

  Query Match             12.4%;  Score 259.6;  DB 3;  Length 2080;
  Best Local Similarity   59.2%;  Pred. No. 1.6e-48;
  Matches  442;  Conservative    0;  Mismatches  304;  Indels    0;  Gaps    0;

Qy        230 CCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATC 289
              | |||||| || |||||||||||||| |    |||   ||   | ||||| ||||||||
Db        112 CAGAGGAGATGCACCGGCTCACGGAAAATGTCTATAAGACCATCATGGAGCAGTTCAACC 171

Qy        290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGA 349
              || | || || || || ||   | ||||||| ||||| ||||| ||  |   | | |||
Db        172 CTAGCCTCCGGAACTTCATCGCCATGGGGAAGAATTACGAGAAGGCACTGGCAGGTGTGA 231

Qy        350 TCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGT 409
                 ||||  |||| |||||  || |   |||  ||||| || ||| | ||||  | |
Db        232 CGTATGCAGCCAAAGGCTACTTTGACGCCCTGGTGAAGATGGGGGAGCTGGCCAGCGAGA 291

Qy        410 CCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC 469
              ||  |  || |  ||||| ||| | || ||| |  ||||  |       ||||| | |
Db        292 GCCAGGGCTCCAAAGAACTCGGAGACGTTCTCTTCCAGATGGCTGAAGTCCACAGGCAGA 351

Qy        470 TCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGG 529
              || |  |   || || ||||   | ||    || ||||| ||| |  |      ||||||
Db        352 TCCAGAATCAGCTGGAAGAAATGCTGAAGTCTTTTCACAACGAGCTGCTTACGCAGCTGG 411

Qy        530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAG 589
              || ||||| | || || |||   |  ||| |||  ||  | ||||| | |||||| || |
Db        412 AGCAGAAGGTGGAGCTGGACTCCAGGTATCTGAGTGCTGCGCTAAAGAAATACCAGACTG 471

Qy        590 AACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
              | || | ||  |||   ||  |  |||| || |  || |||||| |||||||  |  | |
Db        472 AGCAAAGGAGCAAAGGCGACGCCCTGGACAAGTGTCAGGCTGAGCTGAAGAAGCTTCGGA 531

Qy        650 GGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGG 709
               ||| ||||| || |||   ||  | |  || ||     |||| ||  |  |||| | |
Db        532 AGAAGAGCCAGGGCAGCAAGAATCCTCAGAAGTACTCGGACAAGGAGCTGCAGTACATCG 591

Qy        710 AGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGG 769
              |  || | |         ||| | ||  |  |||| | | |  | || || | |||    |
Db        592 ACGCCATCAGCAACAAGCAGGGCGAGCTGGAGAATTACGTGTCCGACGGCTACAAGACCG 651

Qy        770 CTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAA 829
              | |||   || |||  |||||||||||||| ||||| || ||||| || | | | || |
Db        652 CACTGACAGAGGAGTGCAGGCGCTTCTGCTTCCTGGTGGAGAAGCAGTGCGCCGTGGCCA 711
```

```
Qy        830 ACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGT 889
              |  ||          || ||||   ||      || || |||      |||||||| | ||
Db        712 AGAACTCCGCGGCCTACCACTCCAAGGGCAAGGAGCTGCTGCCGCAGAAGCTGCCGCTGT 771

Qy        890 GGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAG 949
              ||||  || ||||||  ||  ||| |||  | || |||       || |  | ||  |
Db        772 GGCAACAGGCCTGTGCCGACCCCAGCAAGATCCCGGAGCGCGCGGTGCAGCTCATGCAGC 831

Qy        950 AAATAAAGACCCCAGCCTCTACCCCC 975
              |  |    | |   | | | |||| |
Db        832 AGGTGGCCAGCAACGGCGCCACCCTC 857


RESULT 3
US-09-046-572-2
; Sequence 2, Application US/09046572
; Patent No. 6589935
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Baughn, Mariah
;    TITLE OF INVENTION:  INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/046,572
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0323-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2080 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT06
;      CLONE:  918158
US-09-046-572-2

  Query Match           12.4%;  Score 259.6;  DB 3;  Length 2080;
  Best Local Similarity  59.2%;  Pred. No. 1.6e-48;
  Matches  442;  Conservative    0;  Mismatches  304;  Indels    0;  Gaps    0;

Qy        230 CCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATC 289
              | |||||| || |||||||||||||| |    |||   ||   | ||||| ||||||| |
Db        112 CAGAGGAGATGCACCGGCTCACGGAAAATGTCTATAAGACCATCATGGAGCAGTTCAACC 171

Qy        290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGA 349
              || | || || || || ||   | ||||||| ||||| ||||| ||  |    | | |||
Db        172 CTAGCCTCCGGAACTTCATCGCCATGGGGAAGAATTACGAGAAGGCACTGGCAGGTGTGA 231

Qy        350 TCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGT 409
              ||||  |||| ||||| || |   |||  ||||| || ||| | ||||  | |
Db        232 CGTATGCAGCCAAAGGCTACTTTGACGCCCTGGTGAAGATGGGGGAGCTGGCCAGCGAGA 291

Qy        410 CCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC 469
              ||  |  || |  ||||| ||| | || ||| |  ||||  |      |||||  | |
Db        292 GCCAGGGCTCCAAAGAACTCGGAGACGTTCTCTTCCAGATGGCTGAAGTCCACAGGCAGA 351

Qy        470 TCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGG 529
              || |  |    || || ||||  | ||  || ||||| ||| |  |      ||||||
Db        352 TCCAGAATCAGCTGGAAGAAATGCTGAAGTCTTTTCACAACGAGCTGCTTACGCAGCTGG 411

Qy        530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAG 589
              || ||||| | || || |||    | ||| |||  ||  | ||||| | |||||| || |
Db        412 AGCAGAAGGTGGAGCTGGACTCCAGGTATCTGAGTGCTGCGCTAAAGAAATACCAGACTG 471

Qy        590 AACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
              | || | ||  |||   ||  |  |||| || |  || |||||| ||||||| |  | |
Db        472 AGCAAAGGAGCAAAGGCGACGCCCTGGACAAGTGTCAGGCTGAGCTGAAGAAGCTTCGGA 531

Qy        650 GGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGG 709
              ||| ||||| || |||   ||  | |  || ||     |||| ||  |  ||||  | |
Db        532 AGAAGAGCCAGGGCAGCAAGAATCCTCAGAAGTACTCGGACAAGGAGCTGCAGTACATCG 591

Qy        710 AGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGG 769
              |  || | |         ||| | ||  |  |||| | | |  | || || | |||    |
Db        592 ACGCCATCAGCAACAAGCAGGGCGAGCTGGAGAATTACGTGTCCGACGGCTACAAGACCG 651

Qy        770 CTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAA 829
              | |||   || |||   |||||||||||||| ||||| || ||||| || | | | || |
Db        652 CACTGACAGAGGAGTGCAGGCGCTTCTGCTTCCTGGTGGAGAAGCAGTGCGCCGTGGCCA 711

Qy        830 ACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGT 889
              |  ||          || ||||   ||      || || |||      |||||||| | ||
Db        712 AGAACTCCGCGGCCTACCACTCCAAGGGCAAGGAGCTGCTGCCGCAGAAGCTGCCGCTGT 771

Qy        890 GGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAG 949
              ||||  || ||||||  ||  ||| |||  | || |||       || |  | ||  |
Db        772 GGCAACAGGCCTGTGCCGACCCCAGCAAGATCCCGGAGCGCGCGGTGCAGCTCATGCAGC 831

Qy        950 AAATAAAGACCCCAGCCTCTACCCCC 975
              |  |    | |   | | | |||| |
Db        832 AGGTGGCCAGCAACGGCGCCACCCTC 857


RESULT 4
US-09-046-572-4
; Sequence 4, Application US/09046572
; Patent No. 6589935
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Baughn, Mariah
;    TITLE OF INVENTION:  INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/046,572
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0323-1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO: 4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2328 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  ADRETUT05
;      CLONE:  2493150
US-09-046-572-4

  Query Match             11.5%;  Score 241.6;  DB 3;  Length 2328;
  Best Local Similarity   58.0%;  Pred. No. 1.9e-44;
  Matches  427;  Conservative    0;  Mismatches  309;  Indels    0;  Gaps    0;

Qy        240 GAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCG 299
              ||||   |     |||||      |  |||   | ||||| |||||||| ||| | || ||
Db        278 GAACTTTCGTGGTGAGAGTTGGCAGGGGACCATCATGGAGCAGTTCAACCCTAGCCTCCG 337

Qy        300 AAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGG 359
              || || ||   | |||||| ||||| ||||| ||  |     | | |||     ||||
Db        338 GAACTTCATCGCCATGGGGAAGAATTACGAGAAGGCACTGGCAGGTGTGACGTATGCAGC 397

Qy        360 AAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTC 419
              |||| |||||  || |  |||  ||||| || ||| | ||||  | |  ||  |  ||
Db        398 CAAAGGCTACTTTGACGCCCTGGTGAAGATGGGGGAGCTGGCCAGCGAGAGCCAGGGCTC 457

Qy        420 AACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAG 479
              |  ||||| ||| | ||  ||| |  ||||  |       |||||  | |  ||  |  |
Db        458 CAAAGAACTCGGAGACGTTCTCTTCCAGATGGCTGAAGTCCACAGGCAGATCCAGAATCA 517

Qy        480 TCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGAT 539
              || || ||||   | ||     || ||||| ||| | |      |||||||| ||||| |
Db        518 GCTGGAAGAAATGCTGAAGTCTTTTCACAACGAGCTGCTTACGCAGCTGGAGCAGAAGGT 577

Qy        540 AGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAA 599
              || || |||   |  ||| |||  ||  | || || | |||||| || || || | ||
Db        578 GGAGCTGGACTCCAGGTATCTGAGTGCTGCGCTGAAGAAATACCAGACTGAGCAAAGGAG 637

Qy        600 TAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCA 659
              |||   ||  |  |||| || |  || |||||| |||||||| |  | | ||| |||||
Db        638 CAAAGGCGACGCCCTGGACAAGTGTCAGGCTGAGCTGAAGAAGCTTCGGAAGAAGAGCCA 697

Qy        660 AGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTAC 719
              || |||   ||  | |  || ||      |||| ||  |  ||||  | ||  || | |
Db        698 GGGCAGCAAGAATCCTCAGAAGTACTCGGACAAGGAGCTGCAGTACATCGACGCCATCAG 757

Qy        720 TTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGA 779
                |||.| ||  |  |||| | | |  | || || | |||    || |||   ||
Db        758 CAACAAGCAGGGCGAGCTGGAGAATTACGTGTCCGACGGCTACAAGACCGCACTGACAGA 817

Qy        780 AGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACA 839
              |||   |||||||||||||| ||||| || ||||| || | | | || ||  ||
Db        818 GGAGCGCAGGCGCTTCTGCTTCCTGGTGGAGAAGCAGTGCGCCGTGGCCAAGAACTCCGC 877

Qy        840 TTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGAC 899
              || ||||   ||      || || |||      |||||||| | ||||||  || |
Db        878 GGCCTACCACTCCAAGGGCAAGGAGCTGCTGCCGCAGAAGCTGCCGCTGTGGCAACAGGC 937

Qy        900 CTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGAC 959
              |||||  ||  ||| |||  | || |||        || |  | ||  |  |  |   |
Db        938 CTGTGCCGACCCCAGCAAGATCCCGGAGCGCGCGGTGCAGCTCATGCAGCAGGTGGCCAG 997

Qy        960 CCCAGCCTCTACCCCC 975

                |   | | | |||| |
Db        998 CAACGGCGCCACCCTC 1013


RESULT 5
US-09-907-794A-376
; Sequence 376, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION:  Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE:  2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
```

; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-376

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
Db 869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy 2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
Db 929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 6
US-09-905-125A-376
; Sequence 376, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-125A-376

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
Db 869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy 2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
Db 929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 7
US-09-902-775A-376
; Sequence 376, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-902-775A-376

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
Db 869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy 2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
Db 929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 8
US-09-906-700-376
; Sequence 376, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,700
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-906-700-376

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
Db 869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy 2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
Db 929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 9
US-09-903-603A-376
; Sequence 376, Application US/09903603A
; Patent No. 6767995
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: GNE.1618P2C12
; CURRENT APPLICATION NUMBER: US/09/903,603A
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-903-603A-376

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;

Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
             ||| |  ||  || |||| |  | | | |   || | | ||| ||| || ||||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986
```

RESULT 10
US-09-904-920A-376
; Sequence 376, Application US/09904920A
; Patent No. 6806352
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,920A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095

```
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-904-920A-376

  Query Match             4.0%;  Score 82.8;  DB 3;  Length 997;
  Best Local Similarity   81.4%;  Pred. No. 7.4e-09;
  Matches   96;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
              |||  |  ||  || ||||  |  | | |    || | | ||| ||| || ||||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986


RESULT 11
US-09-909-064-376
; Sequence 376, Application US/09909064
; Patent No. 6818449
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,064
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-909-064-376

  Query Match             4.0%;  Score 82.8;  DB 3;  Length 997;
  Best Local Similarity   81.4%;  Pred. No. 7.4e-09;
  Matches   96;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
              |||  |  ||  || ||||  |  | | |    || | | ||| ||| || ||||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986


RESULT 12
US-09-905-381A-376
; Sequence 376, Application US/09905381A
; Patent No. 6818746
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
```

```
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,381A
;  CURRENT FILING DATE:  2001-07-13
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-905-381A-376

  Query Match             4.0%;  Score 82.8;  DB 3;  Length 997;
  Best Local Similarity   81.4%;  Pred. No. 7.4e-09;
  Matches   96;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
             |||  |  ||  || |||| |  | | | |   || | | ||| ||| || |||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986


RESULT 13
US-09-906-618-376
; Sequence 376, Application US/09906618
; Patent No. 6828146
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/906,618
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-906-618-376

  Query Match             4.0%;  Score 82.8;  DB 3;  Length 997;
  Best Local Similarity   81.4%;  Pred. No. 7.4e-09;
  Matches   96;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
             |||  |  ||  || |||| |  | | | |   || | | ||| ||| || |||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986
```

RESULT 14
US-09-906-646-376
; Sequence 376, Application US/09906646
; Patent No. 6852848
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,646
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-906-646-376

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
             ||| |  ||  || |||| |  | | | |   || | | ||| ||| || |||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986
```

RESULT 15
US-09-904-462-376
; Sequence 376, Application US/09904462
; Patent No. 6878807
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,462
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

```
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-904-462-376

  Query Match             4.0%;  Score 82.8;  DB 3;  Length 997;
  Best Local Similarity   81.4%;  Pred. No. 7.4e-09;
  Matches   96;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
              |||  |  ||  || |||| |  | | | |    || | | ||| ||| || ||||||||||
Db        869 ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAA 928

Qy       2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986


Search completed: March 20, 2006, 16:06:09
Job time : 382 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:40:03 ; Search time 1705 Seconds
(without alignments)
10165.757 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.....aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2096	100.0	2096	9	US-10-801-292-1	Sequence 1, Appli
	2	2025.2	96.6	2527	9	US-10-756-149-4599	Sequence 4599, Ap
	3	1514.6	72.3	2471	7	US-10-112-944-39	Sequence 39, Appl
	4	1448.4	69.1	1510	7	US-10-112-944-523	Sequence 523, App
	5	1216.4	58.0	2120	6	US-10-172-118-1805	Sequence 1805, Ap
	6	1216.4	58.0	2120	7	US-10-342-887-1805	Sequence 1805, Ap
	7	1071	51.1	1729	9	US-10-450-763-9761	Sequence 9761, Ap
	8	1021.2	48.7	1993	9	US-10-450-763-9762	Sequence 9762, Ap
	9	594	28.3	625	5	US-10-066-543-481	Sequence 481, App
	10	576.4	27.5	1242	3	US-09-925-301-269	Sequence 269, App
	11	576.2	27.5	641	5	US-10-066-543-990	Sequence 990, App
	12	556	26.5	559	3	US-09-878-134-252	Sequence 252, App
	13	520.8	24.8	539	5	US-10-066-543-321	Sequence 321, App
	14	488	23.3	511	3	US-09-960-253-83	Sequence 83, Appl
	15	470.4	22.4	512	3	US-09-960-253-90	Sequence 90, Appl
	16	329	15.7	449	9	US-10-450-763-9759	Sequence 9759, Ap
	17	274.8	13.1	439	9	US-10-450-763-9760	Sequence 9760, Ap
	18	259.6	12.4	2080	6	US-10-435-324-2	Sequence 2, Appli
c	19	255.4	12.2	295	3	US-09-867-701-2669	Sequence 2669, Ap
	20	241.6	11.5	2328	6	US-10-435-324-4	Sequence 4, Appli
c	21	240.2	11.5	45698	3	US-09-984-429-344	Sequence 344, App
c	22	227.4	10.8	1415	9	US-10-970-760-1	Sequence 1, Appli
	23	216.4	10.3	16181	3	US-09-764-847-1426	Sequence 1426, Ap
c	24	216.4	10.3	16181	3	US-09-764-891-6956	Sequence 6956, Ap
	25	216.4	10.3	16181	5	US-10-092-154-1426	Sequence 1426, Ap
c	26	209.6	10.0	476	3	US-09-964-824A-78	Sequence 78, Appl
c	27	209.6	10.0	476	3	US-09-964-824A-491	Sequence 491, App
c	28	209.6	10.0	476	3	US-09-967-768A-205	Sequence 205, App
c	29	209.6	10.0	476	9	US-10-843-641A-5381	Sequence 5381, Ap
c	30	209.6	10.0	476	9	US-10-843-641A-5794	Sequence 5794, Ap
c	31	209.6	10.0	476	9	US-10-843-641A-6350	Sequence 6350, Ap
	32	203.8	9.7	363	6	US-10-313-669-23	Sequence 23, Appl
	33	199	9.5	277	9	US-10-779-543-6174	Sequence 6174, Ap
c	34	187	8.9	454	3	US-09-764-891-820	Sequence 820, App
c	35	184	8.8	490	6	US-10-313-669-24	Sequence 24, Appl
	36	176	8.4	176	3	US-09-867-701-2765	Sequence 2765, Ap
	37	158.4	7.6	2316	6	US-10-094-749-1190	Sequence 1190, Ap
	38	157	7.5	996	3	US-09-876-143-1286	Sequence 1286, Ap
c	39	152	7.3	2242	5	US-10-027-632-102199	Sequence 102199,
c	40	152	7.3	2242	5	US-10-027-632-102200	Sequence 102200,
c	41	152	7.3	2242	5	US-10-027-632-102201	Sequence 102201,
c	42	152	7.3	2242	6	US-10-027-632-102199	Sequence 102199,
c	43	152	7.3	2242	6	US-10-027-632-102200	Sequence 102200,
c	44	152	7.3	2242	6	US-10-027-632-102201	Sequence 102201,
c	45	147.2	7.0	640	4	US-09-925-065A-730687	Sequence 730687,

ALIGNMENTS

RESULT 1
US-10-801-292-1
; Sequence 1, Application US/10801292
; Publication No. US20050202448A1
; GENERAL INFORMATION:
; APPLICANT: LEE, YI-CHAO
; APPLICANT: YUEN, PUI-YEE
; APPLICANT: HUANG, YI-HUEI
; APPLICANT: WU, HUI-CHUAN
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF CANCER
; FILE REFERENCE: 5422-2
; CURRENT APPLICATION NUMBER: US/10/801,292
; CURRENT FILING DATE: 2004-03-15
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-801-292-1

Query Match 100.0%; Score 2096; DB 9; Length 2096;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60

Qy         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120

Qy        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180

Qy        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240

Qy        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
```

```
Qy       301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360

Qy       361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Qy       421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480

Qy       481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540

Qy       541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600

Qy       601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660

Qy       661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720

Qy       721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780

Qy       781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840

Qy       841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900

Qy       901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960

Qy       961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020

Qy      1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080

Qy      1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140

Qy      1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200

Qy      1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260

Qy      1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320

Qy      1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380

Qy      1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Qy      1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500

Qy      1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560

Qy      1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620

Qy      1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680

Qy      1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740

Qy      1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800

Qy      1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Qy      1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920

Qy      1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Qy      1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAA 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAA 2040

Qy      2041 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
```

RESULT 2
US-10-756-149-4599
; Sequence 4599, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AI
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4599
; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2508)..(2508)
; OTHER INFORMATION: n is a, c, g, or t
US-10-756-149-4599

Query Match 96.6%; Score 2025.2; DB 9; Length 2527;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2027; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60

Qy         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120

Qy        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180

Qy        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240

Qy        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300

Qy        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360

Qy        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Qy        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480

Qy        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540

Qy        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600

Qy        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660

Qy        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720

Qy        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780

Qy        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840

Qy        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900

Qy        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960

Qy        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020

Qy       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080

Qy       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140

Qy       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200

Qy       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260

Qy       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320

Qy       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380

Qy       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Qy       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500

Qy       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCCTATCCATGGGGGCAGCT 1560

Qy       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620

Qy       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680

Qy       1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCAGCACCC 1740

Qy       1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800

Qy       1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Qy       1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920

Qy       1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Qy       1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
```

RESULT 3
US-10-112-944-39
; Sequence 39, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom

```
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Zhiwei
;  TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
;  TITLE OF INVENTION:  Secreted Polypeptides
;  FILE REFERENCE: 805A
;  CURRENT APPLICATION NUMBER: US/10/112,944
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: US 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER:.US 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: US 09/552,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: US 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 924
;  SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 39
;   LENGTH: 2471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (61)..(1251)
US-10-112-944-39

  Query Match             72.3%;  Score 1514.6;  DB 7;  Length 2471;
  Best Local Similarity   99.7%;  Pred. No. 1.9e-306;
  Matches 1528;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        499 AAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATAT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATAT 60

Qy        559 ATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAG 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAG 120

Qy        619 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTC 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTC 180

Qy        679 AAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATC 240

Qy        739 CAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 300

Qy        799 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCT 360

Qy        859 GCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 GCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAA 420

Qy        919 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTG 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTG 480

Qy        979 TCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTAC 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTAC 540

Qy       1039 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACC 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACC 600

Qy       1099 AGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 660

Qy       1159 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACG 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACG 720

Qy       1219 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCC 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCC 780

Qy       1279 AACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAG 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAG 840

Qy       1339 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCG 1398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCG 900

Qy       1399 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGC 1458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGC 960

Qy       1459 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1020

Qy       1519 CCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGAT 1578
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        1021 CCCCCACCCGACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGAT 1080

Qy       1579 TCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCT 1638
              |||| |||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        1081 TCGGTCAGGACGACATCCACCTTTAAGGCCCCAGTGTTCAAGCCCGAGACCGCGGCTCCT 1140

Qy       1639 AACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACT 1698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 AACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACT 1200

Qy       1699 GTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAC 1758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAC 1260

Qy       1759 AGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGT 1818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGT 1320

Qy       1819 CTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGC 1878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 CTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGC 1380

Qy       1879 TTAAGCAAATCATG-CTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1937
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TTAAGCAAATCATGCCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1440

Qy       1938 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1500
```

```
Qy       1998 TTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
              ||||||||||||||||||||||||||||||||||
Db       1501 TTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 1533
```

RESULT 4
US-10-112-944-523
; Sequence 523, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 523
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-112-944-523

Query Match 69.1%; Score 1448.4; DB 7; Length 1510;
Best Local Similarity 99.2%; Pred. No. 1.1e-292;
Matches 1498; Conservative 0; Mismatches 6; Indels 6; Gaps 4;

```
Qy        217 ATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATG 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATG 60

Qy        277 GAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCT 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCT 120

Qy        337 GTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAG 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAG 180

Qy        397 ATTGCCACTGGGTCCCCCGTGTCAAC-TGAACTGGGACATGTCCTCATAGAGATTTCAAG 455
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 ATTGCCACTGGGTCCCCCGTGTCAACTTGAACTGGGACATGTCCTCATAGAGATTTCAAG 240

Qy        456 TACCCACAAGAAAC-TCAACGAGAGTCTT---GATGAAAATTTTAAAAAATTCCACAAAG 511
              ||||||||||||| |||||||||||||   ||| ||||||||||||||||||||||||
Db        241 TACCCACAAGAAACTTCAACGAGAGTCTTTGATGGAAAATTTTTAAAAAATTCCACAAAG 300

Qy        512 AGATTATCCATGAGCTGGAGAAGAAGATAGAACTT-GACGTGAAATATATGAACGCAACT 570
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        301 AGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGGACGTGAAATATATGAACGCAACT 360

Qy        571 CTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCT 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCT 420

Qy        631 GAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACAC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACAC 480

Qy        691 AAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATT 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATT 540

Qy        751 GCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGAT 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGAT 600

Qy        811 AAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTG 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTG 660

Qy        871 AATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAA 930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAA 720

Qy        931 ATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCT 990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCT 780

Qy        991 CAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCT 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCT 840

Qy       1051 AAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATC 1110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATC 900

Qy       1111 GATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACA 1170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACA 960

Qy       1171 GGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATG 1230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATG 1020

Qy       1231 ATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTA 1290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTA 1080

Qy       1291 CTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGG 1350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGG 1140

Qy       1351 CTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAG 1410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAG 1200

Qy       1411 TTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTG 1470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTG 1260

Qy       1471 AGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGAC 1530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGAC 1320

Qy       1531 TACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACG 1590
              ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1321 TACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACG 1380

Qy       1591 ACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAAC 1650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAAC 1440

Qy       1651 GGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGC 1710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGC 1500

Qy       1711 CCGACTGTGA 1720
              ||||||||||
Db       1501 CCGACTGTGA 1510
```

RESULT 5
US-10-172-118-1805
; Sequence 1805, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1805
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_018842
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1805

Query Match 58.0%; Score 1216.4; DB 6; Length 2120;
Best Local Similarity 98.3%; Pred. No. 4.4e-244;
Matches 1229; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy        232 GAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCT 291
              | || |||||| |  | |    |          || ||||||||||||||||||||||||
Db        292 GTGGGGGTGAAACTCCCCTGTAATTAGGATCGCCAGAATGTTATGGAACAGTTCAATCCT 351

Qy        292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 411

Qy        352 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 471

Qy        412 CCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTC 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 CCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTC 531

Qy        472 AACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 AACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 591

Qy        532 AAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAA 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 AAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAA 651

Qy        592 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGG 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGG 711

Qy        652 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAG 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAG 771

Qy        712 ACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCT 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCT 831

Qy        772 CTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAAC 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAAC 891

Qy        832 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 951

Qy        892 CAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAA 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 CAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAA 1011

Qy        952 ATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAG 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 ATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAG 1071

Qy       1012 AGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCC 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1072 AGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCC 1131

Qy       1072 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1191

Qy       1132 ACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGT 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 ACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGT 1251

Qy       1192 TTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAG 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 TTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAG 1311

Qy       1252 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1371

Qy       1312 GTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTG 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTG 1431

Qy       1372 TCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACA 1431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACA 1491

Qy       1432 GAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAG 1481
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1492 GAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCTTCAG 1541
```

RESULT 6
US-10-342-887-1805
; Sequence 1805, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene

```
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1805
;   LENGTH: 2120
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-1805

  Query Match             58.0%;  Score 1216.4;  DB 7;  Length 2120;
  Best Local Similarity   98.3%;  Pred. No. 4.4e-244;
  Matches 1229;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy        232 GAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCT 291
              | || |||||| | | |     |          || ||||||||||||||||||||||||||
Db        292 GTGGGGGTGAAACTCCCCTGTAATTAGGATCGCCAGAATGTTATGGAACAGTTCAATCCT 351

Qy        292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 411

Qy        352 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 471

Qy        412 CCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTC 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 CCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTC 531

Qy        472 AACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 AACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 591

Qy        532 AAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAA 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 AAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAA 651

Qy        592 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGG 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGG 711

Qy        652 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAG 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAG 771

Qy        712 ACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCT 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCT 831

Qy        772 CTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAAC 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAAC 891

Qy        832 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 951

Qy        892 CAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAA 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 CAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAA 1011

Qy        952 ATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAG 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 ATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAG 1071

Qy       1012 AGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCC 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1072 AGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCC 1131

Qy       1072 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1191

Qy       1132 ACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGT 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 ACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGT 1251

Qy       1192 TTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAG 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 TTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAG 1311

Qy       1252 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1371

Qy       1312 GTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTG 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTG 1431

Qy       1372 TCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACA 1431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACA 1491

Qy       1432 GAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAG 1481
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1492 GAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCTTCAG 1541

RESULT 7
US-10-450-763-9761
; Sequence 9761, Application US/10450763
; Publication No. US20050196754A1.
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 9761
;   LENGTH: 1729
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (319)..(939)
;   OTHER INFORMATION: 67% homologous to Homo sapiens insulin receptor tyrosine
;   OTHER INFORMATION: kinase substrate,accession number AF119666,Smith-Waterman Score=
;   OTHER INFORMATION: 619.
US-10-450-763-9761

  Query Match             51.1%;  Score 1071;  DB 9;  Length 1729;
  Best Local Similarity   95.1%;  Pred. No. 1.1e-213;
  Matches 1205;  Conservative    0;  Mismatches   45;  Indels   17;  Gaps    9;

Qy        232 GAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCT 291
              | || |||||| | | |     |          || ||||||||||||||||||||||||||
Db        277 GTGGGGGTGAAACTCCCCTGTAATTAGGATCGCCAGAATGTTATGGAACAGTTCAATCCT 336

Qy        292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        337 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 396
Qy        352 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 411
Db        397 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 456
Qy        412 CCCGTGTCAAC-TGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC- 469
Db        457 CCCGTGTCAACTTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACT 516
Qy        470 TCAACGAGAGTCTT---GATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGC 526
Db        517 TCAACGAGAGTCTTTGATGGAAAATTTTTAAAAAATTCCACAAAGAGATTATCCATGAGC 576
Qy        527 TGGAGAAGAAGATAGAACTT-GACGTGAAATATATGAACGCAACTCTAAAAAGATACCAA 585
Db        577 TGGAGAAGAAGATAGAACTTGGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAA 636
Qy        586 ACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATC 645
Db        637 ACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATC 696
Qy        646 AGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTAT 705
Db        697 AGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTAT 756
Qy        706 GTGGAGACCGTTACTTCTCGTCAGAGTGAAATCC-AGAAATTCATTGCAGATGGTTGCAA 764
Db        757 GTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCTAGAAATCCATTGCAGATGGTTGCAA 816
Qy        765 AGAGGCTCTGCTTGAAGAGAAGAGGCG------CTTCTGCTTTCTGGTTGATAAGCACTG 818
Db        817 AGAGGCTCTGCTTGAAAGAGAAAGAGGCGCTTCCTGCTTTCCTGGGTTGAATAAGCACTG 876
Qy        819 TGGCTTTGCAAA--CCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATT-C 875
Db        877 TGGCTTTGCAAAACCCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 936
Qy        876 CAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTG-CCAGAGAAAATCA 934
Db        937 CAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCCAGAGAAAATCA 996
Qy        935 TGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGG 994
Db        997 TGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGG 1056
Qy        995 CTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAAT 1054
Db       1057 CTTCACCCATGATCGAGAGAAGCAATGTGGTAAGGAAAGATTACGACACCCTTTCTAAAT 1116
Qy       1055 GCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATA 1114
Db       1117 GCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATA 1176
Qy       1115 TGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTA 1174
Db       1177 TGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTA 1236
Qy       1175 CTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGA 1234
Db       1237 CTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGA 1296
Qy       1235 AGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCA 1294
Db       1297 AGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCA 1356
Qy       1295 GCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCT 1354
Db       1357 GCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCT 1416
Qy       1355 ATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGC 1414
Db       1417 ATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGC 1476
Qy       1415 TGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAA 1474
Db       1477 TGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAA 1536
Qy       1475 GCATCAG 1481
Db       1537 GCTTCAG 1543
```

RESULT 8
US-10-450-763-9762
; Sequence 9762, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 9762
; LENGTH: 1993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (76)..(501)
; OTHER INFORMATION: 79% homologous to Homo sapiens insulin receptor tyrosine
; OTHER INFORMATION: kinase substrate,accession number AF119666,Smith-Waterman Score=
; OTHER INFORMATION: 578.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1993)
; OTHER INFORMATION: n = a,t,c or g
US-10-450-763-9762

Query Match 48.7%; Score 1021.2; DB 9; Length 1993;
Best Local Similarity 89.3%; Pred. No. 3e-203;
Matches 1211; Conservative 0; Mismatches 38; Indels 107; Gaps 6;

```
Qy         680 AATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCC 739
Db         158 AACTTGACGTGAAATATATGAACTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCC 217
Qy         740 AGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCT 799
Db         218 AGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCT 277
Qy         800 TTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTG 859
Db         278 TTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTG 337
Qy         860 CAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAG 919
Db         338 CAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAG 397
Qy         920 TGCCAGAGAAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGT 979
Db         398 TGCCAGAGAAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGT 457
Qy         980 CTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACG 1039
Db         458 CTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTG---------------- 501
Qy        1040 ACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCA 1099
```

```
Db        502 ------------------------------TCCCTGTTCCTTCTGG-------------- 517

Qy       1100 GTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAA 1159

Db        518 ----------------------------------------GCAATGATCAAAGCGAGG 535

Qy       1160 ATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGG 1219

Db        536 AAATAGGACAGGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGG 595

Qy       1220 GACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCA 1279

Db        596 GACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCA 655

Qy       1280 ACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGA 1339

Db        656 ACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGA 715

Qy       1340 AGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGT 1399

Db        716 AGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGT 775

Qy       1400 CGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCC 1459

Db        776 CGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCC 835

Qy       1460 CCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCC 1519

Db        836 CCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCC 895

Qy       1520 CCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATT 1579

Db        896 CCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATT 955

Qy       1580 CGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTA 1639

Db        956 CGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTA 1015

Qy       1640 ACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTG 1699

Db       1016 ACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTG 1075

Qy       1700 TGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACA 1759

Db       1076 TGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACA 1135

Qy       1760 GCCAAGGACTCTCCC-GGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGC--GCATCCT 1816

Db       I136 GCCAAGGACTCTCCCAGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCCGCATCCCT 1195

Qy       1817 GTCTGCCACGTGCTGACGGTC-GGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCC 1875

Db       1196 GTCTGCCACGTGCTGACGGTCGGGGAAGCTTCAGTGGAAAGGCCTAACTCTAATGTCGCC 1255

Qy       1876 TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC-AAGTTTCTGCCTTT 1934

Db       1256 TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAAGTTTCTGCCTTT 1315

Qy       1935 AAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATT 1994

Db       1316 AAGATAAATGAGTAATAGTCGAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATT 1375

Qy       1995 CTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030

Db       1376 CTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 1411
```

RESULT 9
US-10-066-543-481
; Sequence 481, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 481
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-481

Query Match 28.3%; Score 594; DB 5; Length 625;
Best Local Similarity 100.0%; Pred. No. 4.3e-114;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264

Db         32 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 91

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324

Db         92 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 151

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384

Db        152 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 211

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444

Db        212 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 271

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504

Db        272 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 331

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564

Db        332 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 391

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624

Db        392 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 451

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 684

Db        452 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 511

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744

Db        512 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 571

Qy        745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798

Db        572 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 625
```

RESULT 10
US-09-925-301-269
; Sequence 269, Application US/09925301
; Patent No. US20020052308A1

```
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 269
;   LENGTH: 1242
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (31)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (46)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (460)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1233)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-269

  Query Match             27.5%;  Score 576.4;  DB 3;  Length 1242;
  Best Local Similarity   95.3%;  Pred. No. 2.8e-110;
  Matches  592;  Conservative    1;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       1410 GTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGT 1469
             | || ||  | |  ||   |||   || ||| || |          ||||||||||||||
Db        239 GGTGGTGAGACATTATCTAACACGTGCTGTGTCCTTTTGGGTTTGAGCCCCACACCAGT 298

Qy       1470 GAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGA 1529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGA 358

Qy       1530 CTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGAC 1589
             |||||||||||||||| |||||||||||:||||||||||||||||||||||||||||||
Db        359 CTACTTGGAATGCTTATCCATGGGGGCAGYTGCCGACAGGAGAGCAGATTCGGCCAGGAC 418

Qy       1590 GACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAA 1649
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        419 GACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGNGGCTCCTAACGATGCCAA 478

Qy       1650 CGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCG 1709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 CGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCG 538

Qy       1710 CCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACT 1769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACT 598

Qy       1770 CTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGC 1829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 CTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGC 658

Qy       1830 TGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATC 1889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATC 718

Qy       1890 ATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAA 1949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 ATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAA 778

Qy       1950 TAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAG 2009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         779 TAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAG 838

Qy       2010 TAAACTTGGTTTCAATCTTTA 2030
             |||||||||||||||||||||
Db         839 TAAACTTGGTTTCAATCTTTA 859


RESULT 11
US-10-066-543-990
; Sequence 990, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 990
;   LENGTH: 641
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 5, 362, 424, 427, 444, 449, 565, 567, 584, 602, 609, 611,
;   LOCATION: 624, 629
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-990

  Query Match             27.5%;  Score 576.2;  DB 5;  Length 641;
  Best Local Similarity   97.3%;  Pred. No. 2.3e-110;
  Matches  578;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 107

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 167

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 227

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 287

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 347

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        348 CACAAAGAGATTATNCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 407

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
```

```
                |||||||| ||||||| || |||||||||||||||| |||| ||||||||||||||||||
Db        408 GCAACTCTTAAAAGATNCCNAACAGAACACAAGAATNAATTNGAGTCTTTGGAGAAATCC 467

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 684
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        468 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCCAGGAAGCCGAAACGCACTCAAATAT 527

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744
              |||||||||||||||||||||||||||||||||| | |||||||||||||||| |||
Db        528 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTNTNGTCAGAGTGAAATCCANAAA 587

Qy        745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              ||||||||||||| |||||| | |||||||||||| |||| | ||||||||||
Db        588 TTCATTGCAGATGGNTGCAAANANGCTCTGCTTGAANAGAANAAGCGCTTCTGC 641


RESULT 12
US-09-878-134-252
; Sequence 252, Application US/09878134
; Publication No. US20020086303A1
; GENERAL INFORMATION:
;  APPLICANT: Meagher, Madeleine Joy
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.532
;  CURRENT APPLICATION NUMBER: US/09/878,134
;  CURRENT FILING DATE:  2001-06-07
;  NUMBER OF SEQ ID NOS: 377
;  SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 252
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(559)
;   OTHER INFORMATION: n = A,T,C or G
US-09-878-134-252

  Query Match             26.5%;  Score 556;  DB 3;  Length 559;
  Best Local Similarity   99.5%;  Pred. No. 3.7e-106;
  Matches  556;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        863 AACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGC 922
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          1 AACTACTGAATTCCAAGCTGCCTCGGNGGCAGGAGACCTGTGTTGATGCCATCAAAGTGC 60

Qy        923 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTG 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTG 120

Qy        983 GAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACA 1042
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAACTCCTCANNCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACA 180

Qy       1043 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTC 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTC 240

Qy       1103 CCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATA 300

Qy       1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 360

Qy       1223 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 420

Qy       1283 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 480

Qy       1343 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 540

Qy       1403 ACACGAAGTTGCTGGAAGA 1421
              ||||||||||||||||||||
Db        541 ACACGAAGTTGCTGGAAGA 559


RESULT 13
US-10-066-543-321
; Sequence 321, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 321
;   LENGTH: 539
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 20, 23, 51, 61, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-321

  Query Match             24.8%;  Score 520.8;  DB 5;  Length 539;
  Best Local Similarity   98.7%;  Pred. No. 8.4e-99;
  Matches  522;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||| || |||||||||||||||||||||||||  |.||||||| |||||||||
Db         11 GCGGCCGCANCCNTGTCCCGGGGGCCCGAGGAGGTGAACANGGTCACGGANAGCACCTAC 70

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 NGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 130

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 190

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 250

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 310

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        311 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 370

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 430

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 490

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 733
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 539


RESULT 14
US-09-960-253-83
; Sequence 83, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Lodes, Michael J.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.556
;  CURRENT APPLICATION NUMBER: US/09/960,253
;  CURRENT FILING DATE:  2001-09-20
;  NUMBER OF SEQ ID NOS: 187
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 83
;   LENGTH: 511
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-960-253-83

  Query Match             23.3%;  Score 488;  DB 3;  Length 511;
  Best Local Similarity   99.6%;  Pred. No. 6e-92;
  Matches  510;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy        806 TTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAAC 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAAC 60

Qy        866 TACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAG 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAG 120

Qy        926 AGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAA 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAA 180

Qy        986 CT-CCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACC 1044
              || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        181 CTCCCTCAGGCTTCACCCATGATCGAGAGAAGC-ATGTGGTTAGGAAAGATTACGACACC 239

Qy       1045 CTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCC 1104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCC 299

Qy       1105 TTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAAT 1164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAAT 359

Qy       1165 TCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTG 1224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 TCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTG 419

Qy       1225 AACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAG 1284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAG 479

Qy       1285 ACCTTACTCAGCTTTGCACAGGGAGATGTCAT 1316
              ||||||||||||||||||||||||||||||||
Db        480 ACCTTACTCAGCTTTGCACAGGGAGATGTCAT 511


RESULT 15
US-09-960-253-90
; Sequence 90, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Lodes, Michael J.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.556
;  CURRENT APPLICATION NUMBER: US/09/960,253
;  CURRENT FILING DATE:  2001-09-20
;  NUMBER OF SEQ ID NOS: 187
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 90
;   LENGTH: 512
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-960-253-90

  Query Match             22.4%;  Score 470.4;  DB 3;  Length 512;
  Best Local Similarity   99.8%;  Pred. No. 2.9e-88;
  Matches  471;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         41 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGCGCTGCCGCAAGTTTTGACATTTTCGCAG 100

Qy         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 160

Qy        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 220

Qy        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGGCCCGAGGAGGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 280

Qy        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 340

Qy        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 400

Qy        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 460

Qy        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCA 512


Search completed: March 20, 2006, 14:08:39
Job time : 1708 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:39:54 ; Search time 854 Seconds
(without alignments)
16357.372 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.....aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2028.4	96.8	3617	13	ACN38684	Acn38684 Tumour-as
2	1821.4	86.9	2317	4	AAF32674	Aaf32674 Human cDN
3	1514.6	72.3	2471	12	ADM86946	Adm86946 Human pro
4	1448.4	69.1	1510	12	ADM87430	Adm87430 Human EST
5	1216.4	58.0	2120	13	ADR25944	Adr25944 Breast ca
6	1206.4	57.6	1221	14	ADV43099	Adv43099 Human psy
7	1071	51.1	1729	5	AAS73957	Aas73957 DNA encod
8	1021.2	48.7	1993	5	AAS73958	Aas73958 DNA encod
9	594	28.3	625	11	ADT94962	Adt94962 Colon can
10	594	28.3	625	11	ADX41444	Adx41444 Human cDN
11	576.4	27.5	1242	3	AAC77875	Aac77875 Human can
12	576.2	27.5	641	11	ADT95471	Adt95471 Colon can
13	576.2	27.5	641	11	ADX41953	Adx41953 Human cDN
14	556	26.5	559	6	ABK29726	Abk29726 Colon ade
15	520.8	24.8	539	11	ADT94802	Adt94802 Colon can
16	520.8	24.8	539	11	ADX41284	Adx41284 Human cDN
17	488	23.3	511	6	ABK70212	Abk70212 Human lun
18	470.4	22.4	512	6	ABK70219	Abk70219 Human lun
19	355	16.9	3128	12	ADQ64787	Adq64787 Novel hum
20	329	15.7	449	5	AAS73955	Aas73955 DNA encod
21	274.8	13.1	439	5	AAS73956	Aas73956 DNA encod
22	259.6	12.4	2080	2	AAV72220	Aav72220 Human IRS
23	259.6	12.4	2080	10	ADD35681	Add35681 Human ins
24	259.6	12.4	2080	10	ADE86152	Ade86152 Human IRS
25	258	12.3	2129	13	ADR66181	Adr66181 Human pro
26	258	12.3	2129	14	ADX07028	Adx07028 Cyclin-de
27	258	12.3	2137	2	AAZ52899	Aaz52899 Human pro
28	258	12.3	2877	14	ADX07032	Adx07032 Cyclin-de
29	258	12.3	3168	14	ADX07030	Adx07030 Cyclin-de
30	258	12.3	3255	11	ADL22621	Adl22621 Human dis
31	258	12.3	3458	8	ABX34756	Abx34756 Human mdd
c 32	255.4	12.2	295	6	ABL79691	Abl79691 Human ova
33	241.6	11.5	2328	10	ADD35683	Add35683 Human ins
34	241.6	11.5	2328	10	ADE86154	Ade86154 Human IRS
c 35	240.2	11.5	45698	12	ADJ12490	Adj12490 DNA fragm
c 36	227.4	10.8	1415	14	ADZ77086	Adz77086 IRSp53 ri
37	219.6	10.5	2133	4	AAH18159	Aah18159 Human cDN
38	216.4	10.3	16181	4	ABK42539	Abk42539 Genomic s
c 39	216.4	10.3	16181	4	AAL04268	Aal04268 Human rep
40	216.4	10.3	16181	9	ADB60695	Adb60695 Connectiv
c 41	209.6	10.0	476	6	ABL67044	Abl67044 Thyroid c
c 42	209.6	10.0	476	6	ABL68013	Abl68013 Ovary can
c 43	209.6	10.0	476	6	ABL67457	Abl67457 Thyroid c
44	203.8	9.7	363	10	ADF66284	Adf66284 Human mic
45	203.8	9.7	363	13	ACF87869	Acf87869 Human SIR

ALIGNMENTS

RESULT 1
ACN38684
ID ACN38684 standard; cDNA; 3617 BP.
XX
AC ACN38684;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA324988, SEQ ID NO:2351.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 2351; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)

CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 3617 BP; 941 A; 941 C; 924 G; 811 T; 0 U; 0 Other;

Query Match 96.8%; Score 2028.4; DB 13; Length 3617;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2029; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAG 60
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGCGGCGCTGCCGCAAGTTTTGACATTTTCGCAG 60

Qy         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGC 120

Qy        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCG 180

Qy        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTG 240

Qy        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGA 300

Qy        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 360

Qy        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Qy        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480

Qy        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATA 540

Qy        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAAT 600

Qy        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAA 660

Qy        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACT 720

Qy        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAA 780

Qy        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT 840

Qy        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACC 900

Qy        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC 960

Qy        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAT 1020

Qy       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080

Qy       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCC 1140

Qy       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGA 1200

Qy       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260

Qy       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACG 1320

Qy       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCG 1380

Qy       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Qy       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAG 1500

Qy       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCT 1560

Qy       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAG 1620

Qy       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGA 1680

Qy       1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740

Qy       1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGA 1800
```

```
Qy       1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Qy       1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC 1920

Qy       1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Qy       1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
```

RESULT 2
AAF32674
ID AAF32674 standard; cDNA; 2317 BP.
XX
AC AAF32674;
XX
DT 22-MAR-2001 (first entry)
XX
DE Human cDNA encoding intracellular signalling molecule INTRA37.
XX
KW Human; intracellular signalling molecule; INTRA; immunosuppressive;
KW cytostatic; neuroprotective; nootropic; antiarteriosclerotic; cancer;
KW antiinflammatory; anti-HIV; neuroleptic; antibacterial; antifungal;
KW antiviral; antiparasitic; antihelminthic; antiparkinsonian; AIDS;
KW cell proliferative disorder; arteriosclerosis autoimmune; epilepsy;
KW inflammatory disorder; Addison's disease; gastrointestinal disorder;
KW neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;
KW mental disorder; schizophrenia; anxiety; ss.
XX
OS Homo sapiens.
XX
PN WO200077040-A2.
XX
PD 21-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US016636.
XX
PR 16-JUN-1999; 99US-0139566P.
PR 17-AUG-1999; 99US-0149640P.
PR 09-NOV-1999; 99US-0164417P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Hillman JL, Lal P, Bandman O, Baughn MR;
PI Azimzai Y, Yang J, Reddy R, Lu DAM;
XX
DR WPI; 2001-025334/03.
DR P-PSDB; AAB64405.
XX
PT New human intracellular signaling molecules, useful for the diagnosis,
PT prevention and treatment of cell proliferative, autoimmune, inflammatory,
PT neurological, gastrointestinal, reproductive and developmental disorders.
XX
PS Claim 5; Page 182-183; 192pp; English.
XX
CC Sequences AAF32638 - AAF32689 represent cDNA encoding human intracellular
CC signalling molecules INTRA1 - INTRA52, represented in AAB64369 -
CC AAB64420. Modulators of the intracellular signalling molecules of the
CC invention exhibit immunosuppressive; cytostatic; neuroprotective;
CC nootropic; antiarteriosclerotic; antiinflammatory; anti-HIV; neuroleptic;
CC antibacterial; antifungal; antiviral; antiparasitic; antihelminthic; and
CC antiparkinsonian activity. INTRA polypeptides their agonists and
CC antagonists are useful for the treatment of a condition associated with
CC decreased or increased expression of functional INTRA. Disorders
CC associated with abnormal INTRA expression or activity include cell
CC proliferative disorders e.g. arteriosclerosis and cancers; autoimmune or
CC inflammatory disorders e.g. Addison's disease and acquired
CC immunodeficiency syndrome (AIDS); viral, bacterial, fungal, parasitic,
CC protozoal and helminthic infections; gastrointestinal disorders e.g.
CC dysphagia and irritable bowel syndrome; neurological disorders e.g.
CC epilepsy and Parkinson's disease; prion diseases e.g. Creutzfeldt-Jakob
CC disease and mental disorders e.g. anxiety, schizophrenia and Tourette's
CC disorder. Antibodies immuno specific for the INTRA proteins may also be
CC useful in the diagnosis of the above disorders
XX
SQ Sequence 2317 BP; 676 A; 554 C; 545 G; 542 T; 0 U; 0 Other;

Query Match 86.9%; Score 1821.4; DB 4; Length 2317;
Best Local Similarity 99.9%; Pred. No. 2.3e-304;
Matches 1822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        208 GCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGG 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 GCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGG 89

Qy        268 AATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTAT 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 AATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTAT 149

Qy        328 GAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAG 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAG 209

Qy        388 ATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAG 269

Qy        448 ATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 ATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 329

Qy        508 AAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCA 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 AAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCA 389

Qy        568 ACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAA 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 ACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAA 449

Qy        628 GCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 GCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAA 509

Qy        688 CACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 CACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 569

Qy        748 ATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTT 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 ATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTT 629

Qy        808 GATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTA 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 GATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTA 689

Qy        868 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAG 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAG 749

Qy        928 AAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACT 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 AAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACT 809

Qy        988 CCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTT 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 CCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTT 869

Qy       1048 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTG 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        870 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTG 929

Qy       1108 ATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCA 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 ATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCA 989

Qy       1168 ACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAAC 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        990 ACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAAC 1049

Qy       1228 ATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACC 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1050 ATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACC 1109

Qy       1288 TTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1110 TTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1169

Qy       1348 TGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACG 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1170 TGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACG 1229

Qy       1408 AAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCA 1467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1230 AAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCA 1289

Qy       1468 GTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCC 1527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1290 GTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCC 1349

Qy       1528 GACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGG 1587
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1350 GACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGG 1409

Qy       1588 ACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCC 1647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1410 ACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCC 1469

Qy       1648 AACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTC 1707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1470 AACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTC 1529

Qy       1708 CGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGA 1767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1530 CGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGA 1589

Qy       1768 CTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGT 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1590 CTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGT 1649

Qy       1828 GCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAA 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1650 GCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAA 1709

Qy       1888 TCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1710 TCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGT 1769

Qy       1948 AATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAAC 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1770 AATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAAC 1829

Qy       2008 AGTAAACTTGGTTTCAATCTTTA 2030
              |||||||||||||||||||||||
Db       1830 AGTAAACTTGGTTTCAATCTTTA 1852


RESULT 3
ADM86946
ID   ADM86946 standard; cDNA; 2471 BP.
XX
AC   ADM86946;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Human protein encoding cDNA SEQ ID NO:39.
XX
KW   respiratory; cytostatic; antiarthritic; antiinflammatory;
KW   gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
KW   antirheumatic; gene therapy; molecular weight marker; chromosome marker;
KW   chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
KW   inflammatory condition; arthritis; inflammatory bowel disease;
KW   Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
KW   graft versus host disease; human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004009834-A2.
XX
PD   29-JAN-2004.
XX
PF   19-JUL-2002; 2002WO-US022858.
XX
PR   21-JUL-2001; 2001US-0306971P.
PR   28-MAR-2002; 2002US-00112944.
XX
PA   (NUVE-) NUVELO INC.
XX
PI   Tang YT,  Yang Y,  Weng G,  Zhang J,  Ren F,  Xue A,  Wang J;
PI   Wehrman T,  Ghosh MJ,  Wang D,  Zhao QA,  Wang Z;
XX
DR   WPI; 2004-143291/14.
DR   P-PSDB; ADM87190.
XX
PT   New isolated polynucleotides and polypeptides, useful for treating, e.g.
PT   cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
PT   Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
PT   versus host disease.
XX
PS   Claim 1; SEQ ID NO 39; 591pp; English.
XX
CC   The present invention describes an isolated polynucleotide (I): (a)
CC   comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
CC   which encodes a polypeptide with biological activity, where the
CC   polynucleotide hybridises to (I) under stringent hybridisation conditions
CC   or has greater than 99% sequence identity with (I). (I) has respiratory,
CC   cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
CC   antibacterial, immunosuppressive, antidiabetic and antirheumatic
CC   activities, and can be used in gene therapy. (I) can be used for
CC   generating polynucleotides encoding chimeric or fusion proteins and
CC   heterologous protein sequences. The polynucleotides can be used to
CC   express recombinant protein for analysis, characterisation or therapeutic
CC   use; as markers for tissues in which the corresponding protein is
CC   preferentially expressed; as molecular weight markers on gels; as
CC   chromosome markers or tags to identify chromosomes or to map related gene
CC   positions; to compare with endogenous DNA sequences in patients to
CC   identify potential genetic disorders; as probes to hybridise and discover
CC   genes, related DNA sequences; as a source of information to derive PCR
CC   primers for genetic fingerprinting; as a probe to subtract-out known
CC   sequences in the process of discovering other novel polynucleotides; for
CC   selecting and making oligomers for attachment to a gene chip or other
CC   support, including for examination of expression patterns; to raise anti-
CC   protein antibodies using DNA immunisation techniques; and as an antigen
CC   to raise anti-DNA antibodies or elicit another immune response. The
CC   polynucleotides and polypeptides can also be used as nutritional sources
CC   or supplements, e.g. as a protein or amino acid supplement, as a carbon
CC   source, as a nitrogen source or as a source of carbohydrates. The
CC   polynucleotides and polypeptides can also be used treat cancer. The
CC   compositions are useful for promoting better or faster closure of non-
CC   healing wounds, for the generation and regeneration of tissues, for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues, and conditions resulting from
CC   systemic cytokine damage. The compositions can also be used to treat
CC   inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
CC   Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
CC   or graft versus host disease. The present sequence represents a novel
```

```
CC   human polynucleotide sequence from the present invention. N.B. The
CC   sequences for this patent were obtained from the USPTO web site from an
CC   equivalent US patent US20040048249A1.
XX
SQ   Sequence 2471 BP; 681 A; 634 C; 585 G; 571 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 1514.6;  DB 12;  Length 2471;
  Best Local Similarity   99.7%;  Pred. No. 1.6e-251;
  Matches 1528;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        499 AAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATAT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATAT 60

Qy        559 ATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAG 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAG 120

Qy        619 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTC 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTC 180

Qy        679 AAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATC 240

Qy        739 CAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 300

Qy        799 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCT 360

Qy        859 GCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAA 420

Qy        919 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTG 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTG 480

Qy        979 TCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTAC 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTAC 540

Qy       1039 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACC 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACC 600

Qy       1099 AGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 660

Qy       1159 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACG 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACG 720

Qy       1219 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCC 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCC 780

Qy       1279 AACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAG 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAG 840

Qy       1339 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCG 1398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCG 900

Qy       1399 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGC 1458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGC 960

Qy       1459 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1020

Qy       1519 CCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGAT 1578
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       1021 CCCCCACCCGACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGAT 1080

Qy       1579 TCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCT 1638
              |||| |||||||||||||||||||||||||| || |||||||||||||||||||||||||
Db       1081 TCGGTCAGGACGACATCCACCTTTAAGGCCCCAGTGTTCAAGCCCGAGACCGCGGCTCCT 1140

Qy       1639 AACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACT 1698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACT 1200

Qy       1699 GTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAC 1758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAC 1260

Qy       1759 AGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGT 1818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGT 1320

Qy       1819 CTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGC 1878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGC 1380

Qy       1879 TTAAGCAAATCATG-CTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1937
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTAAGCAAATCATGCCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1440

Qy       1938 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1500

Qy       1998 TTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
              |||||||||||||||||||||||||||||||||
Db       1501 TTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 1533


RESULT 4
ADM87430
ID   ADM87430 standard; cDNA; 1510 BP.
XX
AC   ADM87430;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Human EST derived nucleotide sequence SEQ ID NO:523.
XX
KW   respiratory; cytostatic; antiarthritic; antiinflammatory;
KW   gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
KW   antirheumatic; gene therapy; molecular weight marker; chromosome marker;
KW   chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
KW   inflammatory condition; arthritis; inflammatory bowel disease;
KW   Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
KW   graft versus host disease; human; expressed sequence tag; EST; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004009834-A2.
XX
PD   29-JAN-2004.
XX
PF   19-JUL-2002; 2002WO-US022858.
XX
PR   21-JUL-2001; 2001US-0306971P.
PR   28-MAR-2002; 2002US-00112944.
XX
PA   (NUVE-) NUVELO INC.
XX
```

PI Tang YT, Yang Y, Weng G, Zhang J, Ren F, Xue A, Wang J;
PI Wehrman T, Ghosh MJ, Wang D, Zhao QA, Wang Z;
XX
DR WPI; 2004-143291/14.
DR P-PSDB; ADM87648.
XX
PT New isolated polynucleotides and polypeptides, useful for treating, e.g.
PT cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
PT Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
PT versus host disease.
XX
PS Example 2; SEQ ID NO 523; 591pp; English.
XX
CC The present invention describes an isolated polynucleotide (I): (a)
CC comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
CC which encodes a polypeptide with biological activity, where the
CC polynucleotide hybridises to (I) under stringent hybridisation conditions
CC or has greater than 99% sequence identity with (I). (I) has respiratory,
CC cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
CC antibacterial, immunosuppressive, antidiabetic and antirheumatic
CC activities, and can be used in gene therapy. (I) can be used for
CC generating polynucleotides encoding chimeric or fusion proteins and
CC heterologous protein sequences. The polynucleotides can be used to
CC express recombinant protein for analysis, characterisation or therapeutic
CC use; as markers for tissues in which the corresponding protein is
CC preferentially expressed; as molecular weight markers on gels; as
CC chromosome markers or tags to identify chromosomes or to map related gene
CC positions; to compare with endogenous DNA sequences in patients to
CC identify potential genetic disorders; as probes to hybridise and discover
CC genes, related DNA sequences; as a source of information to derive PCR
CC primers for genetic fingerprinting; as a probe to subtract-out known
CC sequences in the process of discovering other novel polynucleotides; for
CC selecting and making oligomers for attachment to a gene chip or other
CC support, including for examination of expression patterns; to raise anti-
CC protein antibodies using DNA immunisation techniques; and as an antigen
CC to raise anti-DNA antibodies or elicit another immune response. The
CC polynucleotides and polypeptides can also be used as nutritional sources
CC or supplements, e.g. as a protein or amino acid supplement, as a carbon
CC source, as a nitrogen source or as a source of carbohydrates. The
CC polynucleotides and polypeptides can also be used treat cancer. The
CC compositions are useful for promoting better or faster closure of non-
CC healing wounds, for the generation and regeneration of tissues, for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, and conditions resulting from
CC systemic cytokine damage. The compositions can also be used to treat
CC inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
CC Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
CC or graft versus host disease. The present sequence represents an
CC expressed sequence tag (EST) derived nucleotide sequence from the present
CC invention. N.B. The sequences for this patent were obtained from the
CC USPTO web site from an equivalent US patent US20040048249A1.
XX
SQ Sequence 1510 BP; 462 A; 372 C; 366 G; 310 T; 0 U; 0 Other;

Query Match 69.1%; Score 1448.4; DB 12; Length 1510;
Best Local Similarity 99.2%; Pred. No. 3.9e-240;
Matches 1498; Conservative 0; Mismatches 6; Indels 6; Gaps 4;

```
Qy      217 ATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATG 276
Db        1 ATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATG 60

Qy      277 GAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCT 336
Db       61 GAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCT 120

Qy      337 GTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAG 396
Db      121 GTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAG 180

Qy      397 ATTGCCACTGGGTCCCCCGTGTCAAC-TGAACTGGGACATGTCCTCATAGAGATTTCAAG 455
Db      181 ATTGCCACTGGGTCCCCCGTGTCAACTTGAACTGGGACATGTCCTCATAGAGATTTCAAG 240

Qy      456 TACCCACAAGAAAC-TCAACGAGAGTCTT---GATGAAAATTTTAAAAAATTCCACAAAG 511
Db      241 TACCCACAAGAAACTTCAACGAGAGTCTTTGATGGAAAATTTTTAAAAAATTCCACAAAG 300

Qy      512 AGATTATCCATGAGCTGGAGAAGAAGATAGAACTT-GACGTGAAATATATGAACGCAACT 570
Db      301 AGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGGACGTGAAATATATGAACGCAACT 360

Qy      571 CTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCT 630
Db      361 CTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCT 420

Qy      631 GAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACAC 690
Db      421 GAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACAC 480

Qy      691 AAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATT 750
Db      481 AAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATT 540

Qy      751 GCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGAT 810
Db      541 GCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGAT 600

Qy      811 AAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTG 870
Db      601 AAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTG 660

Qy      871 AATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAA 930
Db      661 AATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAA 720

Qy      931 ATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCT 990
Db      721 ATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCT 780

Qy      991 CAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCT 1050
Db      781 CAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCT 840

Qy     1051 AAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATC 1110
Db      841 AAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATC 900

Qy     1111 GATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACA 1170
Db      901 GATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACA 960

Qy     1171 GGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATG 1230
Db      961 GGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATG 1020

Qy     1231 ATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTA 1290
Db     1021 ATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTA 1080

Qy     1291 CTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGG 1350
Db     1081 CTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGG 1140

Qy     1351 CTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAG 1410
Db     1141 CTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAG 1200

Qy     1411 TTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTG 1470
Db     1201 TTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGGCAAGCCCCACACCAGTG 1260

Qy     1471 AGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGAC 1530
Db     1261 AGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGAC 1320
```

```
Qy       1531 TACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACG 1590
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACG 1380

Qy       1591 ACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAAC 1650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAAC 1440

Qy       1651 GGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGC 1710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGC 1500

Qy       1711 CCGACTGTGA 1720
             ||||||||||
Db       1501 CCGACTGTGA 1510
```

RESULT 5
ADR25944
ID ADR25944 standard; DNA; 2120 BP.
XX
AC ADR25944;
XX
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #1805.
XX
KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004WO-US001100.
XX
PR 15-JAN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
XX
DR WPI; 2004-593473/57.
XX
PT Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 1805; 226pp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
SQ Sequence 2120 BP; 625 A; 513 C; 464 G; 518 T; 0 U; 0 Other;

```
  Query Match             58.0%;  Score 1216.4;  DB 13;  Length 2120;
  Best Local Similarity   98.3%;  Pred. No. 3.7e-200;
  Matches 1229;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy        232 GAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCT 291
              | || |||||| | | |    |          || ||||||||||||||||||||||||
Db        292 GTGGGGGTGAAACTCCCCTGTAATTAGGATCGCCAGAATGTTATGGAACAGTTCAATCCT 351

Qy        292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 411

Qy        352 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 471

Qy        412 CCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTC 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 CCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTC 531

Qy        472 AACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 AACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 591

Qy        532 AAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAA 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 AAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAA 651

Qy        592 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGG 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGG 711

Qy        652 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAG 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAG 771

Qy        712 ACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCT 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCT 831

Qy        772 CTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAAC 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAAC 891

Qy        832 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 951

Qy        892 CAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAA 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 CAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAA 1011

Qy        952 ATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAG 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 ATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAG 1071

Qy       1012 AGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCC 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1072 AGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCC 1131

Qy       1072 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1191

Qy       1132 ACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGT 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 ACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGT 1251

Qy       1192 TTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAG 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 TTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAG 1311

Qy       1252 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1371

Qy       1312 GTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTG 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTG 1431

Qy       1372 TCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACA 1431
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACA 1491

Qy       1432 GAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAG 1481
                 ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1492 GAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCTTCAG 1541


RESULT 6
ADV43099
ID   ADV43099 standard; cDNA; 1221 BP.
XX
AC   ADV43099;
XX
DT   10-MAR-2005  (first entry)
XX
DE   Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 727.
XX
KW   microarray; psychoneuroendocrinimmune; chronic fatigue;
KW   non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW   cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004108899-A2.
XX
PD   16-DEC-2004.
XX
PF   04-JUN-2004; 2004WO-US017686.
XX
PR   04-JUN-2003; 2003US-0475915P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nicholson A,  Vernon SD;
XX
DR   WPI; 2005-031682/03.
XX
PT   New microarray comprising probes for genes involved in
PT   psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT   condition associated with PNI activity, e.g., inflammatory or infectious
PT   diseases.
XX
PS   Claim 1; SEQ ID NO 727; 254pp; English.
XX
CC   The invention relates to a new microarray which comprises probes for
CC   genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC   microarray is useful in diagnosing a condition associated with PNI
CC   activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC   cancer and infection. The present sequence represents a
CC   psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC   specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
CC   SEQ ID NO 1829 are provided.
XX
SQ   Sequence 1221 BP; 397 A; 279 C; 288 G; 257 T; 0 U; 0 Other;

  Query Match             57.6%;  Score 1206.4;  DB 14;  Length 1221;
  Best Local Similarity   99.9%;  Pred. No. 1.9e-198;
  Matches 1207;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         274 ATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAA 333
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAA 60

Qy         334 GCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGT 393
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGT 120

Qy         394 GAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCA 453
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCA 180

Qy         454 AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAG 513
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAG 240

Qy         514 ATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTA 573
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTA 300

Qy         574 AAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 633
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 360

Qy         634 TTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAA 693
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAA 420

Qy         694 GAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCA 753
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCA 480

Qy         754 GATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAG 813
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAG 540

Qy         814 CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 873
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 600

Qy         874 TCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATC 933
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATC 660

Qy         934 ATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAG 993
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 ATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAG 720

Qy         994 GCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA 1053
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA 780

Qy        1054 TGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 1113
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 840

Qy        1114 ATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGT 1173
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 ATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGT 900

Qy        1174 ACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATG 1233
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 ACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATG 960

Qy        1234 AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC 1293
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC 1020

Qy        1294 AGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTC 1353
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTC 1080

Qy        1354 TATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTG 1413
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTG 1140

Qy        1414 CTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGA 1473
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGA 1200

Qy        1474 AGCATCAG 1481
               ||| ||||
Db        1201 AGCTTCAG 1208


RESULT 7
```

```
AAS73957
ID   AAS73957 standard; cDNA; 1729 BP.
XX
AC   AAS73957;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9761.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09770.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 9761; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1729 BP; 517 A; 426 C; 389 G; 397 T; 0 U; 0 Other;

  Query Match           51.1%;  Score 1071;  DB 5;  Length 1729;
  Best Local Similarity   95.1%;  Pred. No. 4.1e-175;
  Matches 1205;  Conservative    0;  Mismatches   45;  Indels   17;  Gaps    9;

Qy        232 GAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAACAGTTCAATCCT 291
              | || |||||| |  | |    |          || ||||||||||||||||||||||||
Db        277 GTGGGGGTGAAACTCCCCTGTAATTAGGATCGCCAGAATGTTATGGAACAGTTCAATCCT 336

Qy        292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 351
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 396

Qy        352 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 411
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 456

Qy        412 CCCGTGTCAAC-TGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAAC- 469
              |||||||||||  |||||||||||||||||||||||||||||||||||||||||||||| |
Db        457 CCCGTGTCAACTTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACT 516

Qy        470 TCAACGAGAGTCTT---GATGAAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGC 526
              |||||||||||||       ||| ||||||||||||||||||||||||||||||||||||
Db        517 TCAACGAGAGTCTTTGATGGAAAATTTTTAAAAAATTCCACAAAGAGATTATCCATGAGC 576

Qy        527 TGGAGAAGAAGATAGAACTT-GACGTGAAATATATGAACGCAACTCTAAAAAGATACCAA 585
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        577 TGGAGAAGAAGATAGAACTTGGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAA 636

Qy        586 ACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATC 645
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 ACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATC 696

Qy        646 AGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTAT 705
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 AGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTAT 756

Qy        706 GTGGAGACCGTTACTTCTCGTCAGAGTGAAATCC-AGAAATTCATTGCAGATGGTTGCAA 764
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        757 GTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCTAGAAATCCATTGCAGATGGTTGCAA 816

Qy        765 AGAGGCTCTGCTTGAAGAGAAGAGGCG------CTTCTGCTTTCTGGTTGATAAGCACTG 818
              |||||||||||||||    | ||  |     || ||    |  | |  ||||||||||
Db        817 AGAGGCTCTGCTTGAAAGAGAAAGAGGCGCTTCCTGCTTTCCTGGGTTGAATAAGCACTG 876

Qy        819 TGGCTTTGCAAA--CCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATT-C 875
              ||||||||||||  |||||||||||||||||||||||||||||||||||||||||| |
Db        877 TGGCTTTGCAAAACCCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 936

Qy        876 CAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTG-CCAGAGAAAATCA 934
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        937 CAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCCAGAGAAAATCA 996

Qy        935 TGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGG 994
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 TGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGG 1056

Qy        995 CTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAAT 1054
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1057 CTTCACCCATGATCGAGAGAAGCAATGTGGTAAGGAAAGATTACGACACCCTTTCTAAAT 1116

Qy       1055 GCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATA 1114
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 GCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATA 1176

Qy       1115 TGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTA 1174
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 TGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTA 1236

Qy       1175 CTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGA 1234
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 CTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGA 1296

Qy       1235 AGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCA 1294
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 AGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCA 1356

Qy       1295 GCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCT 1354
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 GCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCT 1416

Qy       1355 ATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGC 1414
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 ATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGC 1476

Qy       1415 TGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAA 1474
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 TGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAA 1536

Qy       1475 GCATCAG 1481
                || ||||
Db       1537 GCTTCAG 1543


RESULT 8
AAS73958
ID   AAS73958 standard; cDNA; 1993 BP.
XX
AC   AAS73958;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9762.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09771.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 9762; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1993 BP; 542 A; 491 C; 472 G; 487 T; 0 U; 1 Other;

  Query Match             48.7%;  Score 1021.2;  DB 5;  Length 1993;
  Best Local Similarity   89.3%;  Pred. No. 1.6e-166;
  Matches 1211;  Conservative    0;  Mismatches   38;  Indels  107;  Gaps    6;
```

```
Qy        680 AATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCC 739
              ||  |||    ||| | ||  | ||||||||||||||||||||||||||||||||||||
Db        158 AACTTGACGTGAAATATATGAACTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCC 217

Qy        740 AGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCT 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 AGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCT 277

Qy        800 TTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTG 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTG 337

Qy        860 CAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAG 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAG 397

Qy        920 TGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGT 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 TGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGT 457

Qy        980 CTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACG 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTG---------------- 501

Qy       1040 ACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCA 1099
                                              ||| | |||||| ||
Db        502 ----------------------------TCCCTGTTCCTTCTGG-------------- 517

Qy       1100 GTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAA 1159
                                                 |||   ||||| |
Db        518 -----------------------------------------GCAATGATCAAAGCGAGG 535

Qy       1160 ATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGG 1219
              | |    |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 AAATAGGACAGGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGG 595

Qy       1220 GACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCA 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 GACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCA 655

Qy       1280 ACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGA 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        656 ACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGA 715

Qy       1340 AGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGT 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        716 AGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGT 775

Qy       1400 CGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCC 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        776 CGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCC 835

Qy       1460 CCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        836 CCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCC 895

Qy       1520 CCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATT 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        896 CCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATT 955

Qy       1580 CGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTA 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        956 CGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTA 1015

Qy       1640 ACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTG 1699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1016 ACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTG 1075

Qy       1700 TGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACA 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1076 TGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACA 1135
```

```
Qy       1760 GCCAAGGACTCTCCC-GGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGC--GCATCCT 1816
             |||||||||||||||| |||||||||||||||||||||||||||||||||||      |||
Db       1136 GCCAAGGACTCTCCCAGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCCGCATCCCT 1195

Qy       1817 GTCTGCCACGTGCTGACGGTC-GGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCC 1875
             |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db       1196 GTCTGCCACGTGCTGACGGTCGGGGAAGCTTCAGTGGAAAGGCCTAACTCTAATGTCGCC 1255

Qy       1876 TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGAC-AAGTTTCTGCCTTT 1934
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1256 TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAAGTTTCTGCCTTT 1315

Qy       1935 AAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATT 1994
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1316 AAGATAAATGAGTAATAGTCGAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATT 1375

Qy       1995 CTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
             |||||||||||||||||||||||||||||||||||||
Db       1376 CTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 1411
```

RESULT 9
ADT94962
ID ADT94962 standard; cDNA; 625 BP.
XX
AC ADT94962;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #481.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 481; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 625 BP; 215 A; 127 C; 159 G; 124 T; 0 U; 0 Other;

```
  Query Match             28.3%;  Score 594;  DB 11;  Length 625;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-93;
  Matches  594;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 91

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 151

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 211

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 271

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 331

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 391

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 451

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 511

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 571

Qy        745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 625
```

RESULT 10
ADX41444
ID ADX41444 standard; cDNA; 625 BP.
XX
AC ADX41444;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 481.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.

```
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 481; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 215 A; 127 C; 159 G; 124 T; 0 U; 0 Other;

  Query Match             28.3%;  Score 594;  DB 11;  Length 625;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-93;
  Matches  594;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 91

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 151

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 211

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 271

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 331

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 391

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 451

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 511

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 571

Qy        745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 625


RESULT 11
AAC77875
ID   AAC77875 standard; cDNA; 1242 BP.
XX
AC   AAC77875;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:269.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005882.
XX
PR   12-MAR-1999;   99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43666.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer.
XX
PS   Claim 1; Page 832-833; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC   AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
```

```
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1242 BP; 322 A; 291 C; 281 G; 338 T; 0 U; 10 Other;

  Query Match           27.5%;  Score 576.4;  DB 3;  Length 1242;
  Best Local Similarity  95.3%;  Pred. No. 6.2e-90;
  Matches  592;  Conservative    1;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       1410 GTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGT 1469
             | || ||  | |  ||   |||   || ||| || |          ||||||||||||||
Db        239 GGTGGTGAGACATTATCTAACACGTGCTGTGTCCTTTTTGGGTTTGAGCCCCACACCAGT 298

Qy       1470 GAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGA 1529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGA 358

Qy       1530 CTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGAC 1589
             |||||||||||||||| ||||||||||||:||||||||||||||||||||||||||||||
Db        359 CTACTTGGAATGCTTATCCATGGGGGCAGYTGCCGACAGGAGAGCAGATTCGGCCAGGAC 418

Qy       1590 GACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAA 1649
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        419 GACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGNGGCTCCTAACGATGCCAA 478

Qy       1650 CGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCG 1709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 CGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCG 538

Qy       1710 CCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACT 1769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACT 598

Qy       1770 CTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGC 1829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 CTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGC 658

Qy       1830 TGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATC 1889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATC 718

Qy       1890 ATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAA 1949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 ATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAA 778

Qy       1950 TAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAG 2009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 TAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAG 838

Qy       2010 TAAACTTGGTTTCAATCTTTA 2030
             |||||||||||||||||||||
Db        839 TAAACTTGGTTTCAATCTTTA 859


RESULT 12
ADT95471
ID   ADT95471 standard; cDNA; 641 BP.
XX
AC   ADT95471;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #990.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 990; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 641 BP; 217 A; 127 C; 154 G; 129 T; 0 U; 14 Other;

  Query Match           27.5%;  Score 576.2;  DB 11;  Length 641;
  Best Local Similarity  97.3%;  Pred. No. 6.2e-90;
  Matches  578;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 107

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 167

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 227

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 287

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        288 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 347

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        348 CACAAAGAGATTATNCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 407

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
              |||||||| ||||||| || |||||||||||||||| |||| ||||||||||||||||||
Db        408 GCAACTCTTAAAAGATNCCNAACAGAACACAAGAATNAATTNGAGTCTTTGGAGAAATCC 467

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCCAGGAAGCCGAAACGCACTCAAATAT 684
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        468 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCCAGGAAGCCGAAACGCACTCAAATAT 527

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744
              |||||||||||||||||||||||||||||||||| | |||||||||||||||| |||
Db        528 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTNTNGTCAGAGTGAAATCCANAAA 587

Qy        745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              |||||||||||||| |||||| | |||||||||||| |||| | ||||||||||
Db        588 TTCATTGCAGATGGNTGCAAANANGCTCTGCTTGAANAGAANAAGCGCTTCTGC 641
```

RESULT 13
ADX41953
ID ADX41953 standard; cDNA; 641 BP.
XX
AC ADX41953;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 990.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 990; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.

XX
SQ Sequence 641 BP; 217 A; 127 C; 154 G; 129 T; 0 U; 14 Other;

Query Match 27.5%; Score 576.2; DB 11; Length 641;
Best Local Similarity 97.3%; Pred. No. 6.2e-90;
Matches 578; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 107

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 167

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 227

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 287

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 347

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        348 CACAAAGAGATTATNCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 407

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
              |||||||| ||||||| || |||||||||||||||| |||| ||||||||||||||||||
Db        408 GCAACTCTTAAAAGATNCCNAACAGAACACAAGAATNAATTNGAGTCTTTGGAGAAATCC 467

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCCAGGAAGCCGAAACGCACTCAAATAT 684
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        468 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCCAGGAAGCCGAAACGCACTCAAATAT 527

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744
              |||||||||||||||||||||||||||||||||| | |||||||||||||||| |||
Db        528 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTNTNGTCAGAGTGAAATCCANAAA 587

Qy        745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGC 798
              |||||||||||||| |||||| | |||||||||||| |||| | ||||||||||
Db        588 TTCATTGCAGATGGNTGCAAANANGCTCTGCTTGAANAGAANAAGCGCTTCTGC 641
```

RESULT 14
ABK29726
ID ABK29726 standard; cDNA; 559 BP.
XX
AC ABK29726;
XX
DT 23-APR-2002 (first entry)
XX
DE Colon adenocarcinoma-specific cDNA #252.
XX
KW Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196389-A2.
XX
PD 20-DEC-2001.
XX
PF 07-JUN-2001; 2001WO-US018574.
XX
PR 09-JUN-2000; 2000US-0210667P.
PR 22-NOV-2000; 2000US-0252614P.
XX
PA (CORI-) CORIXA CORP.
XX

PI Meagher MJ, King GE, Xu J, Secrist H;
XX
DR WPI; 2002-098052/13.
XX
PT New isolated polynucleotide encoding a polypeptide comprising a portion
PT of colon tumor protein, for detection, diagnosis and therapy of human
PT colon cancer.
XX
PS Claim 1; Page 178-179; 211pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) encoding a
CC polypeptide (II) comprising a portion of a colon tumour protein. A new
CC oligonucleotide (III) that hybridises to (I) is useful for determining
CC the presence of a cancer in a patient. (II) or antigen presenting cells
CC expressing (I) are useful for stimulating and/or expanding T cells
CC specific for a tumour protein, by contacting T cells with (I), (II) or
CC antigen-presenting cells that express (I). (I), (II), or antigen
CC presenting cells that express (II) are useful for treating colon cancer
CC in a patient by incubating CD4+ and/or CD8+ T cells isolated from a
CC patient with (I), (II), or antigen presenting cells that express (II), so
CC that T cells proliferate, and administering to the patient an effective
CC amount of the proliferated T cells, thus inhibiting the development of a
CC cancer in the patient. A new composition is useful for stimulating an
CC immune response in a patient. (I) or (II) is useful in vaccines and
CC pharmaceutical compositions for prevention and treatment of colon cancer
CC and for the diagnosis and monitoring of the cancers. (I), (II) or an
CC antibody against (II) is useful for detection, diagnosis and/ or therapy
CC of human colon cancer. (I) is useful as a probe or primer for nucleic
CC acid hybridisation, and in the design and preparation of ribozyme
CC molecules for inhibiting expression of (II) in tumour cells. ABK29475-
CC ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of
CC the invention
XX
SQ Sequence 559 BP; 160 A; 148 C; 132 G; 116 T; 0 U; 3 Other;

Query Match 26.5%; Score 556; DB 6; Length 559;
Best Local Similarity 99.5%; Pred. No. 1.8e-86;
Matches 556; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        863 AACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGC 922
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db          1 AACTACTGAATTCCAAGCTGCCTCGGNGGCAGGAGACCTGTGTTGATGCCATCAAAGTGC 60

Qy        923 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTG 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTG 120

Qy        983 GAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACA 1042
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAACTCCTCANNCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACA 180

Qy       1043 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTC 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTC 240

Qy       1103 CCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATA 300

Qy       1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGAC 360

Qy       1223 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACA 420

Qy       1283 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGG 480

Qy       1343 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGT 540

Qy       1403 ACACGAAGTTGCTGGAAGA 1421
              ||||||||||||||||||||
Db        541 ACACGAAGTTGCTGGAAGA 559
```

RESULT 15
ADT94802
ID ADT94802 standard; cDNA; 539 BP.
XX
AC ADT94802;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #321.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 321; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX

```
SQ   Sequence 539 BP; 191 A; 106 C; 132 G; 105 T; 0 U; 5 Other;

  Query Match             24.8%;  Score 520.8;  DB 11;  Length 539;
  Best Local Similarity   98.7%;  Pred. No. 2.1e-80;
  Matches  522;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        205 GCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
              ||||||||| || |||||||||||||||||||||||||  | |||||||  |||||||||
Db         11 GCGGCCGCANCCNTGTCCCGGGGGCCCGAGGAGGTGAACANGGTCACGGANAGCACCTAC 70

Qy        265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 NGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAAT 130

Qy        325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCC 190

Qy        385 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 250

Qy        445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 310

Qy        505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 370

Qy        565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCC 430

Qy        625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATAT 490

Qy        685 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 733
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 GAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 539


Search completed: March 20, 2006, 13:54:24
Job time : 861 secs
```

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:39:51 ; Search time 8382 Seconds
(without alignments)
11699.562 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.....aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1475.8	70.4	1479	11	DQ051599	DQ051599 Homo sapi
	2	1223.4	58.4	3245	4	AK004918	AK004918 Mus muscu
	3	1189.2	56.7	2214	4	AK008203	AK008203 Mus muscu
	4	950	45.3	1479	11	DQ051600	DQ051600 Pan trogl
	5	852.2	40.7	909	5	BU153780	BU153780 AGENCOURT
	6	830	39.6	848	5	BU184787	BU184787 AGENCOURT
	7	811.4	38.7	895	5	BQ888698	BQ888698 AGENCOURT
c	8	803.4	38.3	814	6	CA450094	CA450094 UI-CF-FN0
	9	794.8	37.9	877	5	BU184917	BU184917 AGENCOURT
	10	792.2	37.8	916	5	BQ688252	BQ688252 AGENCOURT
	11	775.2	37.0	812	3	BI333892	BI333892 602999431
c	12	768.4	36.7	779	3	BM981661	BM981661 UI-CF-EN1
	13	762.6	36.4	943	5	BU500695	BU500695 AGENCOURT
	14	760.4	36.3	996	3	BI335025	BI335025 602999031
	15	757.2	36.1	933	2	BG761459	BG761459 602718666
	16	751.8	35.9	880	3	BQ233322	BQ233322 AGENCOURT
	17	739	35.3	739	7	CN392816	CN392816 170006002
c	18	732.8	35.0	763	5	BU689777	BU689777 UI-CF-FN0
c	19	724.2	34.6	876	5	BU176351	BU176351 AGENCOURT
	20	722.8	34.5	756	8	DR005158	DR005158 TC113380
	21	718.8	34.3	747	2	BI255527	BI255527 602975670
	22	717	34.2	717	3	BI335902	BI335902 Tmdl19 Hu
	23	713.4	34.0	720	2	BF732309	BF732309 nae09a07.
	24	707.8	33.8	725	2	BG682313	BG682313 602629631
	25	696.8	33.2	919	2	BG674550	BG674550 602620038
c	26	688.2	32.8	712	5	BU633552	BU633552 UI-H-FL1-
	27	686.4	32.7	924	5	BQ922694	BQ922694 AGENCOURT
	28	676	32.3	676	3	BM817746	BM817746 K-EST0084
	29	665.8	31.8	847	2	BG034422	BG034422 602302824
	30	663.4	31.7	690	1	AW473168	AW473168 xq12f05.x
	31	661.8	31.6	690	1	AI924372	AI924372 wn60b08.x
c	32	647.6	30.9	662	3	BM968354	BM968354 UI-CF-DU1
	33	647.4	30.9	989	2	BG033827	BG033827 602301207
	34	638.2	30.4	688	1	AW374090	AW374090 PM4-BT054
	35	637	30.4	648	7	CV027819	CV027819 6171 Full
	36	612.2	29.2	680	2	BE732245	BE732245 601568088
	37	607	29.0	914	2	BG335429	BG335429 602404087
c	38	604.8	28.9	890	5	BU181833	BU181833 AGENCOURT
	39	597.4	28.5	789	3	BI560834	BI560834 603254194
	40	595	28.4	1009	5	BY708155	BY708155 BY708155
	41	593.4	28.3	902	3	BI332789	BI332789 602979635
	42	585.4	27.9	588	2	BG766055	BG766055 602738434
	43	583.8	27.9	863	5	BU501073	BU501073 AGENCOURT
	44	581	27.7	581	3	BM852920	BM852920 K-EST0134
	45	578.6	27.6	931	5	BQ960401	BQ960401 AGENCOURT

ALIGNMENTS

RESULT 1
DQ051599
LOCUS DQ051599 1479 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens LOC55971 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ051599
VERSION DQ051599.1 GI:66904813
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1479)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1479)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES Location/Qualifiers
source 1. .1479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
gene <1. .>1479
/gene="LOC55971"
/locus_tag="HC19618"
ORIGIN

Query Match 70.4%; Score 1475.8; DB 11; Length 1479;
Best Local Similarity 99.9%; Pred. No. 1e-290;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        274 ATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAA 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAA 60

Qy        334 GCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGT 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGT 120

Qy        394 GAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCA 453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCA 180

Qy        454 AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAG 513
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAG 240

Qy        514 ATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTA 573
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTA 300

Qy        574 AAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 360

Qy        634 TTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAA 693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAA 420

Qy        694 GAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCA 753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCA 480

Qy        754 GATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAG 813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAG 540

Qy        814 CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 600

Qy        874 TCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATC 933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATC 660

Qy        934 ATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAG 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAG 720

Qy        994 GCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA 780

Qy       1054 TGCTCACCAAAGATGCCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 1113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGCTCACCAAAGATGCCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 840

Qy       1114 ATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGT 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGT 900

Qy       1174 ACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATG 960

Qy       1234 AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC 1020

Qy       1294 AGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTC 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTC 1080

Qy       1354 TATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTG 1413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTG 1140

Qy       1414 CTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGA 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGA 1200

Qy       1474 AGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTAC 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTAC 1260

Qy       1534 TTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACA 1593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACA 1320

Qy       1594 TCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGG 1653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGG 1380

Qy       1654 ACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCG 1713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1381 ACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGTCCG 1440

Qy       1714 ACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGA 1752
              |||||||||||||||||| ||||||||||||||||||
Db       1441 ACTGTGACGAATGATCGCTCAGCACCCATCATTCGATGA 1479
```

RESULT 2
AK004918

```
LOCUS       AK004918               3245 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male liver cDNA, RIKEN full-length enriched
            library, clone:1300006M19 product:INSULIN RECEPTOR TYROSINE KINASE
            SUBSTRATE homolog [Homo sapiens], full insert sequence.
ACCESSION   AK004918
VERSION     AK004918.2  GI:26338407
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
```

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3245)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT On Dec 10, 2002 this sequence version replaced gi:12836466. Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTTAATTAAACCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.
FEATURES Location/Qualifiers

source 1. .3245
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1300006M19"
/db_xref="taxon:10090"
/clone="1300006M19"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS 273. .1817
/note="unnamed protein product; INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE homolog [Homo sapiens] (SPTR|Q9UHR4, evidence: FASTY, 90.3%ID, 99.2%length, match=1206) putative"
/codon_start=1
/protein_id="BAB23669.1"
/db_xref="GI:12836467"
/translation="MSRGPEEVNRLTENTYRNVMEQFNPGLRNLINLGKNYEKAVNAM ILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNETLDENFKKFHKDIIH ELEKKTELDVKYMNATLKRYQAEHRNKLDSLEKSQAELKKIRRKSQGGRNALKYEHKE IEYVETVTSRQSEIQKFIADGCKEALLEEKRRFCFLVDKHCSFASHIHYYHMQSAELL NSKLPRWQETCCDATKVPEKIMNMIEEIKTPISTPVSGTPQPSPMIERSKMIGKDYDT LSKYSPKMPPAPSVKAYTSPLIDMFNNPATAAQSSEKTNNSTANTGEDPSLQRSVSVA TGLNMMKKQKVKTIFPHTAGNNKTLLSFAQGDVLTLLIPEEKDGWLYGEHDTTKARGW FPSSYTKLLEENEAMSVPTPSPAPVRSISTVDLTEKSSVVIPPPDYLECLSMGATSDK KAGAPKVPSASTFRAPVSRPDATSTSPSDANGTAKPPFLSGENPFATVKLRPTVTNDR SAPIIR"

polyA_signal 3227. .3232
/note="putative"

polyA_site 3245
/note="putative"

ORIGIN

Query Match 58.4%; Score 1223.4; DB 4; Length 3245;
Best Local Similarity 77.9%; Pred. No. 3.9e-239;
Matches 1600; Conservative 0; Mismatches 421; Indels 34; Gaps 9;

```
Qy      24 CAGGGCGGGGCTGCCGCAAGTTTTGACATTTTCGCAGCGGAGACGCGCGCGGGCACTCTC 83
           |||  ||  ||  |  ||||||||||||||||| |||||  |||||||||  | | | |||||
Db      84 CAGACCGAAGCGCCGCAAAGTTTTGACATTTTGGCAGCAGAGACGCGCACTGACTCTCTC 143

Qy      84 GGGCCGACGGCTGCGGCGGCGGCCGACCCTCCAGAGCCCCTTAGTCGCGCCCCGGCCCTC 143
           || ||| |||| |||||||||||||| ||| ||  ||  || |||     | |  | |||
Db     144 CCGCTGACAGCTGTGGCGGCGGCCGACTCTCAAGGACCACTCAGT----CTCAAGTCCTA 199

Qy     144 CCGCTGCCCGGAGTCCGGCGGCCACGAGGCCCAGCCGCGTCCTCCCGCGCTTGCTCGCCC 203
               ||||  | |||||||||||||||||| |||||||||||||||||| || | |||
Db     200 GGCGCTCCCGCCGCCCGGCGGCCACGAGGCCCCGCCGCGTCCTCCCGCGCTCGCCCTCCC 259

Qy     204 GGCGGCCGCAGCCATGTCCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTA 263
           || ||| |||||||||||| |||||||| |||||||||||||||||||||||||| ||| ||
Db     260 GGAGGCTGCAGCCATGTCTCGGGGGCCGGAGGAGGTGAACCGGCTCACGGAGAACACGTA 319

Qy     264 CCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAA 323
           ||||||||| |||||||| ||||| ||||| || || || |||||||||| || |||||
Db     320 CCGGAATGTGATGGAACAATTCAACCCTGGACTACGGAACTTAATAAACCTAGGAAAAAA 379

Qy     324 TTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGC 383
           |||||| ||||| || || ||||||||||||||||||||||||||||||| || |||||
Db     380 TTATGAAAAAGCGGTGAATGCTATGATCCTGGCAGGAAAAGCCTACTACGACGGCGTGGC 439

Qy     384 CAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCAT 443
           |||||| ||||| |||||||||||| || |||||||| || || || ||||||||||||
Db     440 CAAGATTGGTGAAATTGCCACTGGGTCTCCTGTGTCAACAGAGCTAGGCCATGTCCTCAT 499

Qy     444 AGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATT 503
           ||||||||| || |||||||||||||||| |||| |||||||||||||| ||||||||
Db     500 AGAGATTTCCAGCACCCACAAGAAACTCAATGAGACTCTTGATGAAAATTTCAAAAAATT 559

Qy     504 CCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAA 563
           ||| || || |||||||||||||||||||||||||| ||||||||| || || ||||||||
Db     560 CCATAAGGATATTATCCATGAGCTGGAGAAGAAGACAGAACTTGATGTAAAGTATATGAA 619

Qy     564 CGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATC 623
           ||| ||||| |||||||| ||| |||| |||| |||||||||||| ||||||||||||||
Db     620 CGCTACTCTTAAAAGATATCAAGCAGAGCACAGGAATAAATTAGATTCTTTGGAGAAATC 679

Qy     624 CCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATA 683
           || || ||| |||||||||||||||| ||||| |||||   | ||||||||||||||||||
Db     680 TCAGGCGGAGCTGAAGAAGATCAGAAGAAAAAGTCAAGGTGGACGAAACGCACTCAAATA 739

Qy     684 TGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAA 743
           || |||||||||||||||||||||||| || || ||||| |||||||| ||||||||
Db     740 CGAGCACAAAGAAATTGAGTATGTGGAGACGGTGACGTCTCGCCAGAGTGAGATCCAGAA 799

Qy     744 ATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCT 803
           |||||||| ||||| |||||||| || |||||||| || || |||||||||||||||||
Db     800 GTTCATTGCCGATGGCTGCAAAGAAGCCCTGCTTGAGGAAAAAAGGCGCTTCTGCTTTCT 859

Qy     804 GGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGA 863
           ||| || ||||||||  |||||||  | ||||||||  || ||||||  | ||||||||||
Db     860 GGTGGACAAGCACTGCAGCTTTGCCAGTCACATACACTACTATCACATGCAGTCTGCAGA 919

Qy     864 ACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCC 923
           |||| |||||||||||||| |||||||||| ||||||  ||||||||  ||||| ||
Db     920 GTTACTTAATTCCAAGCTGCCCCGGTGGCAGGAAACCTGTTGTGATGCCACAAAAGTCCC 979

Qy     924 AGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGG 983
           |||||||||||||| ||||| ||||| |||||||| ||   ||| ||||||||||| ||
Db     980 AGAGAAAATCATGAACATGATAGAAGAGATAAAGACGCCCATCTCCACCCCCGTGTCCGG 1039

Qy     984 AACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACAC 1043
```

```
Db       1040 GACGCCGCAGCCATCGCCCATGATCGAGAGAAGCAAGATGATTGGGAAAGATTATGACAC 1099

Qy       1044 CCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCC 1103

Db       1100 CCTTTCTAAATACTCACCAAAGATGCCCCCTGCTCCTTCAGTCAAAGCCTACACCAGTCC 1159

Qy       1104 CTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAA 1163

Db       1160 TTTGATTGACATGTTCAATAACCCAGCTACAGCTGCACAGAGTTCAGAAAAAACAAACAA 1219

Qy       1164 TTCAACAG---GTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGG 1220

Db       1220 CTCCACAGCAAACACTGGGGAAGATCCCAGTCTACAGCGATCGGTTTCTGTTGCAACTGG 1279

Qy       1221 ACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAA 1280

Db       1280 GCTGAACATGATGAAGAAGCAGAAGGTGAAGACCATCTTCCCTCACACGGCTGGCAACAA 1339

Qy       1281 CAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAA 1340

Db       1340 TAAGACCCTACTGAGCTTTGCCCAGGGGGACGTCCTCACACTGCTCATCCCTGAGGAGAA 1399

Qy       1341 GGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTC 1400

Db       1400 GGACGGCTGGCTCTACGGGGAGCATGACACCACCAAAGCGAGGGGCTGGTTCCCATCGTC 1459

Qy       1401 GTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCC 1460

Db       1460 ATATACAAAGTTGCTGGAAGAAAAT------GAAGCAATGTCTGTCCCTACACCAAGCCC 1513

Qy       1461 CACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCC 1520

Db       1514 TGCACCCGTGAGAAGCATCAGCACGGTGGACCTGACTGAGAAGAGCAGCGTTGTCATCCC 1573

Qy       1521 CCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTC 1580

Db       1574 ACCTCCCGACTACTTGGAGTGCCTGTCCATGGGAGCCACCTCAGACAAGAAAGCGGGCGC 1633

Qy       1581 GGCCAGG------ACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGA------GAC 1628

Db       1634 TCCCAAGGTGCCTTCTGCTTCTACCTTCAGAGCGCCAGTGTCCAGACCAGACGCCACATC 1693

Qy       1629 CGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCC 1688

Db       1694 CACCTCTCCCAGCGATGCCAACGGGACCGCAAAGCCGCCATTCCTCAGTGGAGAAAACCC 1753

Qy       1689 CTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCG 1748

Db       1754 CTTTGCCACCGTGAAACTCCGGCCAACAGTGACAAATGACCGATCAGCACCGATCATCCG 1813

Qy       1749 ATG-AGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGG 1807

Db       1814 ATGAAGATGCAGGGGACAGTGCCAGTCCCTCTACTCGAGTTCTCACCCACAGGGCGATAG 1873

Qy       1808 GCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTA 1867

Db       1874 GTATA-CATTGCTGGACAGCAGGCACTGTCAAAGATTGCAGAGTGCCTGATGTCAATGAA 1932

Qy       1868 ATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTC 1927

Db       1933 GCCCCACTTGAGCAAATCAATGTACTTTCTCAGATTGACAGAGGTGGGTTTACACATTTC 1992

Qy       1928 TGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCT 1987

Db       1993 TGTCTTTAAGGTAAATGAG---TAGCTTAGTGAATAAATCAATCATTTGAAATATTTTCT 2049

Qy       1988 TCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAA 2047

Db       2050 T----TTCTATTCAAAAAGCAGTAAATTTGGCTCAAATTTTTAAATTTTTATTTTTAAGG 2105

Qy       2048 AAAAAAAAAAAAAA 2062

Db       2106 AAATAGCCAGAATAA 2120
```

RESULT 3
AK008203

LOCUS	AK008203 2214 bp mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus adult male small intestine cDNA, RIKEN full-length enriched library, clone:2010012D17 product:INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE homolog [Homo sapiens], full insert sequence.
ACCESSION	AK008203
VERSION	AK008203.1 GI:12842246
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Carninci,P. and Hayashizaki,Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2214)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,

```
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.jp/) for further
            details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 5.0 and subtraction to
            Rot = 20.0. Second strand cDNA was prepared with the primer adapter
            of sequence[5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES              Location/Qualifiers
     source           1. .2214
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="FANTOM_DB:2010012D17"
                      /db_xref="taxon:10090"
                      /clone="2010012D17"
                      /sex="male"
                      /tissue_type="small intestine"
                      /clone_lib="RIKEN full-length enriched mouse cDNA library"
                      /dev_stage="adult"
     misc_feature     177. .1717
                      /note="INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE homolog
                      [Homo sapiens] (SPTR|Q9UHR4, evidence: FASTY, 90.3%ID,
                      99.2%length, match=1206)
                      putative"
     polyA_signal     2192. .2197
                      /note="putative"
     polyA_site       2214
                      /note="putative"
ORIGIN

  Query Match             56.7%;  Score 1189.2;  DB 4;  Length 2214;
  Best Local Similarity   77.8%;  Pred. No. 3.6e-232;
  Matches 1586;  Conservative    0;  Mismatches  413;  Indels   39;  Gaps   11;

Qy         41 AAGTTTTGACATTTTCGCAGCGGAGACGCGCGCGGGCACTCTCGGGCCGACGGCTGCGGC 100
              ||||||||||||||| ||||| ||||||||| | | | |||||  || ||| |||| |||
Db          5 AAGTTTTGACATTTTGGCAGCAGAGACGCGCACTGACTCTCTCCCGCTGACAGCTGTGGC 64

Qy        101 GGCGGCCGACCCTCCAGAGCCCCTTAGTCGCGCCCCGGCCCTCCCGCTGCCCGGAGTCCG 160
              |||||||||| ||| ||  || || |||    | |  | |||       ||||  | |||
Db         65 GGCGGCCGACTCTCAAGGACCACTCAGT----CTCAAGTCCTAGGCGCTCCCGCCGCCCG 120

Qy        161 GCGGCCACGAGGCCCAGCCGCGTCCTCCCGCGCTTGCTCGCCCGGCGGCCGCAGCCATGT 220
              |||  |||||||||| |||||||||||||||||| || | ||||| ||| ||||||||||
Db        121 GCGCGCACGAGGCCCCGCCGCGTCCTCCCGCGCTCGCCCTCCCGGAGGCTGCAGCCATGT 180

Qy        221 CCCGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAGCACCTACCGGAATGTTATGGAAC 280
              | |||||||| |||||||||||||||||||||||| ||| ||||||||||| |||||||
Db        181 CTCGGGGGCCGGAGGAGGTGAACCGGCTCACGGAGAACACGTACCGGAATGTGATGGAAC 240

Qy        281 AGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAA 340
              | ||||| ||||| || || || |||||||||| || |||||||||| ||||| || |
Db        241 AATTCAACCCTGGACTACGGAACTTAATAAACCTAGGAAAAAATTATGAAAAAGCGGTGA 300

Qy        341 ACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTG 400
              | |||||||||||||||||||||||||||||||| || |||||||     ||||| ||||
Db        301 ATGCTATGATCCTGGCAGGAAAAGCCTACTACGACGGCGTGGCCA----TGGTGAAATTG 356

Qy        401 CCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCC 460
              |||||||||| || |||||||| || || || |||||||||||||||||||| || ||||
Db        357 CCACTGGGTCTCCTGTGTCAACAGAGCTAGGCCATGTCCTCATAGAGATTTCCAGCACCC 416

Qy        461 ACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATTATCC 520
              |||||||||||||| |||| |||||||||||||| |||||||||| || || |||||||
Db        417 ACAAGAAACTCAATGAGACTCTTGATGAAAATTTCAAAAAATTCCATAAGGATATTATCC 476

Qy        521 ATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGAT 580
              |||||||||||||||||| ||||||||| || || |||||||||| ||||| |||||||
Db        477 ATGAGCTGGAGAAGAAGACAGAACTTGATGTAAAGTATATGAACGCTACTCTTAAAAGAT 536

Qy        581 ACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGA 640
              | ||| |||| |||| |||||||||||| |||||||||||||| || || ||| ||||||
Db        537 ATCAAGCAGAGCACAGGAATAAATTAGATTCTTTGGAGAAATCTCAGGCGGAGCTGAAGA 596

Qy        641 AGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTG 700
              |||||||||| ||||| ||||| | |||||||||||||||||| || ||||||||||||
Db        597 AGATCAGAAGAAAAAGTCAAGGTGGACGAAACGCACTCAAATACGAGCACAAAGAAATTG 656

Qy        701 AGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTT 760
              |||||||||||| || || ||||| |||||||| |||||||| |||||||| ||||| |
Db        657 AGTATGTGGAGACGGTGACGTCTCGCCAGAGTGAGATCCAGAAGTTCATTGCCGATGGCT 716

Qy        761 GCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTG 820
              ||||||| || |||||||| || || |||||||||||||||||| || ||||||||
Db        717 GCAAAGAAGCCCTGCTTGAGGAAAAAAGGCGCTTCTGCTTTCTGGTGGACAAGCACTGCA 776

Qy        821 GCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGC 880
              ||||||| |  |||||||| || |||||| | |||||||||||  |||| ||||||||||
Db        777 GCTTTGCCAGTCACATACACTACTATCACATGCAGTCTGCAGAGTTACTTAATTCCAAGC 836

Qy        881 TGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATA 940
              |||| |||||||||| |||||| |||||||  ||||| |||||||||||||||||| |
Db        837 TGCCCCGGTGGCAGGAAACCTGTTGTGATGCCACAAAAGTCCCAGAGAAAATCATGAACA 896

Qy        941 TGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCAC 1000
              |||| ||||| |||||||| ||   ||| |||||||||| || || || ||| | || |
Db        897 TGATAGAAGAGATAAAGACGCCCATCTCCACCCCCGTGTCCGGGACGCCGCAGCCATCGC 956

Qy       1001 CCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCAC 1060
              |||||||||||||||||||  || || |||||||||| ||||||||||||||||| |||||
Db        957 CCATGATCGAGAGAAGCAAGATGATTGGGAAAGATTATGACACCCTTTCTAAATACTCAC 1016

Qy       1061 CAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTA 1120
              |||||||||||| ||||||||||.|| ||| || |||||||| ||||| || ||||| |
Db       1017 CAAAGATGCCCCCTGCTCCTTCAGTCAAAGCCTACACCAGTCCTTTGATTGACATGTTCA 1076

Qy       1121 ATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAG---GTACTT 1177
              |||||||||| || ||||| | || |||| |||     ||| || || ||||     |||
Db       1077 ATAACCCAGCTACAGCTGCACAGAGTTCAGAAAAAACAAACAACTCCACAGCAAACACTG 1136

Qy       1178 CCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGA 1237
              |||||||||||| |||||||||| ||||| |||||||| || ||||||||||||||||
Db       1137 GGGAAGATCCCAGTCTACAGCGATCGGTTTCTGTTGCAACTGGGCTGAACATGATGAAGA 1196

Qy       1238 AGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCT 1297
              ||||||| |||||||||||||||||| ||||| || |||  ||| |||||| |||| ||||
Db       1197 AGCAGAAGGTGAAGACCATCTTCCCTCACACGGCTGGCAACAATAAGACCCTACTGAGCT 1256

Qy       1298 TTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATG 1357
              |||| ||||| || ||| |||| |||||||||| |||||||||| |||||||||| |
Db       1257 TTGCCCAGGGGGACGTCCTCACACTGCTCATCCCTGAGGAGAAGGACGGCTGGCTCTACG 1316

Qy       1358 GAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGG 1417
              | || || |||    |||| |||||||| |||||||| ||||| || || ||||||||||
Db       1317 GGGAGCATGACACCACCAAAGCGAGGGGCTGGTTCCCATCGTCATATACAAAGTTGCTGG 1376

Qy       1418 AAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCA 1477
              ||||||||       |||||| || | || || || ||||||||  |||| ||||||||||
Db       1377 AAGAAAAT------GAAGCAATGTCTGTCCCTACACCAAGCCCTGCACCCGTGAGAAGCA 1430

Qy       1478 TCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGG 1537
              |||||| ||| || || |||||||| ||||| |||||||||||| || ||||||||||||
Db       1431 TCAGCACGGTGGACCTGACTGAGAAGAGCAGCGTTGTCATCCCACCTCCCGACTACTTGG 1490

Qy       1538 AATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGG------ACGA 1591
              | ||| |||||||||| ||  |  | |||| || ||| |   |  ||| |       |
```

```
Db       1491 AGTGCCTGTCCATGGGAGCCACCTCAGACAAGAAAGCGGGCGCTCCCAAGGTGCCTTCTG 1550

Qy       1592 CATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGA------GACCGCGGCTCCTAACGATG 1645

Db       1551 CTTCTACCTTCAGAGCGCCAGTGTCCAGACCAGACGCCACATCCACCTCTCCCAGCGATG 1610

Qy       1646 CCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAAC 1705

Db       1611 CCAACGGGACCGCAAAGCCGCCATTCCTCAGTGGAGAAAACCCCTTTGCCACCGTGAAAC 1670

Qy       1706 TCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATG-AGAGGACAGCCAA 1764

Db       1671 TCCGGCCAACAGTGACAAATGACCGATCAGCACCGATCATCCGATGAAGATGCAGGGGAC 1730

Qy       1765 GGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCA 1824

Db       1731 AGTGCCAGTCCCTCTACTCGAGTTCTCACCCACAGGGCGATAGGTATA-CATTGCTGGAC 1789

Qy       1825 CGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGC 1884

Db       1790 AGCAGGCACTGTCAAAGATTGCAGAGTGCCTGATGTCAATGAAGCCCCACTTGAGCAAAT 1849

Qy       1885 AAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATG 1944

Db       1850 CAATGTACTTTCTCAGATTGACAGAGGTGGGTT-ACACATTTCTGTCTTTAAGGTAAATG 1908

Qy       1945 AGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGA 2004

Db       1909 AG---TAGCTTAGTGAATAAATCAATCATTTGAAATATTTTCTT----TTCTATTCAAAA 1961

Qy       2005 AACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2062

Db       1962 AGCAGTAAATTTGGCTCAAATTTTTAAATTTTTATTTTTAAGGAAATAGCCAGAATAA 2019
```

RESULT 4
DQ051600
LOCUS DQ051600 1479 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes LOC55971 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ051600
VERSION DQ051600.1 GI:66904814
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1479)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1479)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES Location/Qualifiers
source 1. .1479
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1. .>1479
/gene="LOC55971"
/locus_tag="HC19618"
ORIGIN

Query Match 45.3%; Score 950; DB 11; Length 1479;
Best Local Similarity 64.4%; Pred. No. 2.6e-183;
Matches 953; Conservative 0; Mismatches 526; Indels 0; Gaps 0;

```
Qy        274 ATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAA 333

Db          1 ATGGAACANTTCAATCCTGGGCTGCGAAATNNAANNNNCCTGGGGAAAAATTATGAGAAA 60

Qy        334 GCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGT 393

Db         61 GCTGTAAACGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120

Qy        394 GAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCA 453

Db        121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGACATGTCCTCATAGAGATTTCA 180

Qy        454 AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAG 513

Db        181 AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAG 240

Qy        514 ATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTA 573

Db        241 ATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAANGCAACTCTA 300

Qy        574 AAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 633

Db        301 AAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 360

Qy        634 TTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAA 693

Db        361 TTGAAGAAGATCAGAAGGAAAAGCCNNGGAAGCCGAAACGCACTCAAATATGAACACAAA 420

Qy        694 GAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCA 753

Db        421 GAANNNNNNNNNNNNNNNNNGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCA 480

Qy        754 GATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAG 813

Db        481 GATGGTTGCAAAGAGGNTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAG 540

Qy        814 CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 873

Db        541 CACTGTGGCTTTGCAAACCACATACATTATTATNNCTTACAGTCTGCAGAACTACTGAAT 600

Qy        874 TCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATC 933

Db        601 TCNNNGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATC 660

Qy        934 ATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAG 993

Db        661 ATGAATATGATTGAAGAAATAAAGACCCCAGCCTCTACCCCCGTGTCCGGAACTCCTCAG 720

Qy        994 GCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA 1053

Db        721 GCTTCANCCATGATCGAGAGAAGCAATGTGNNNNNNNNNNNNNNNNNNNCCCTTTCTAAA 780

Qy       1054 TGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 1113

Db        781 TGCTCACCAAAGNNNCCCCCTGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATNGNT 840

Qy       1114 ATGTTTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGT 1173

Db        841 ATGTTTAATAACCCAGCCATGGCTGCCCCGNATTCACAAAGGGTAAATAATTCAACAGNN 900

Qy       1174 ACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATG 1233

Db        901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960

Qy       1234 AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC 1293
```

```
Db        961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020

Qy       1294 AGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTC 1353

Db       1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080

Qy       1354 TATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTG 1413

Db       1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140

Qy       1414 CTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGA 1473
                                                                 |||||||||
Db       1141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCAGTGAGA 1200

Qy       1474 AGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTAC 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTAC 1260

Qy       1534 TTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACA 1593
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1261 TTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGAACGACA 1320

Qy       1594 TCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGG 1653
              ||
Db       1321 TCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1380

Qy       1654 ACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCG 1713
                                  ||||||||||||||||||||||||||||||||||||
Db       1381 NNNNNNNNNNNNNNNNNNNNNNCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCG 1440

Qy       1714 ACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGA 1752
              |||||||||||||||||||||||||||||||||||||||
Db       1441 ACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGA 1479
```

RESULT 5
BU153780

```
LOCUS       BU153780                909 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_7780106 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137888
            5', mRNA sequence.
ACCESSION   BU153780
VERSION     BU153780.1  GI:22667312
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13452  row: g  column: 09
            High quality sequence stop: 732.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6137888"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.  Library constructed by Life
                     Technologies."
ORIGIN
```

```
  Query Match             40.7%;  Score 852.2;  DB 5;  Length 909;
  Best Local Similarity   96.8%;  Pred. No. 2.3e-163;
  Matches  880;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy        960 CCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAA 1019
              ||  || ||  ||| || ||| |     |    ||||||||||||||||||||||||||
Db          1 CCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCTTCACCCATGATCGAGAGAAGCAA 60

Qy       1020 TGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCC 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCC 120

Qy       1080 TTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGC 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGC 180

Qy       1140 CCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCG 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCG 240

Qy       1200 ATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTT 1259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTT 300

Qy       1260 CCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCAC 1319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCAC 360

Qy       1320 GCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGC 1379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGC 420

Qy       1380 GAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGT 1439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGT 480

Qy       1440 GACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGA 1499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGA 540

Qy       1500 GAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGC 1559
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        541 GAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTATCCATGGGGGCAGC 600

Qy       1560 TGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAA 1619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAA 660

Qy       1620 GCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGG 1679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGG 720

Qy       1680 AGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACC 1739
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        721 AGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCAGCACC 780

Qy       1740 CATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGG 1799
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        781 CATCATTCGATGAGAGGACAGCCCAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGG 840

Qy       1800 AATGATGGGCGCATCCTGTCTGCCACGTGC-TGACGGTCGGGAAGCTTCAGTGGAGAGGC 1858
              |||||||||||| |||||||||| ||||| |||||   |||| |||||||||||||||
Db        841 AATGATGGGCGCGTCCTGTCTGCCCCGTGCTTGACGTCGGGGAACTTTCAGTGGAGAGGC 900

Qy       1859 CTAACTCTA 1867
              |||||||||
Db        901 CTAACTCTA 909
```

```
RESULT 6
BU184787
LOCUS       BU184787                848 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7825409 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6093297
            5', mRNA sequence.
ACCESSION   BU184787
VERSION     BU184787.1  GI:22698771
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 848)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13362  row: e  column: 10
            High quality sequence stop: 710.
FEATURES             Location/Qualifiers
     source          1. .848
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6093297"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.  Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             39.6%;  Score 830;  DB 5;  Length 848;
  Best Local Similarity   99.8%;  Pred. No. 7.9e-159;
  Matches  841;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy        953 TAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGA 60

Qy       1013 GAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCC 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCC 120

Qy       1073 CCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCA 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCA 180

Qy       1133 CGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTT 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTT 240

Qy       1193 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA 300

Qy       1253 CCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATG 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATG 360

Qy       1313 TCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGT 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGT 420

Qy       1373 CCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAG 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAG 480

Qy       1433 AAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACT 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACT 540

Qy       1493 TGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGG 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGG 600

Qy       1553 GGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAG 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAG 660

Qy       1613 CGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTC 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTC 720

Qy       1673 TCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCT 1732
              |||||||||||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db        721 TCAGCGGAGAAAA-CCCTTTGCCACTGTGANACTCCGCCCGACTGTGACGAATGATCGCT 779

Qy       1733 CGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCC 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCC 839

Qy       1793 CTT 1795
              |||
Db        840 CTT 842


RESULT 7
BQ888698
LOCUS       BQ888698                895 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8074718 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6086238
            5', mRNA sequence.
ACCESSION   BQ888698
VERSION     BQ888698.1  GI:22280712
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2320  row: o  column: 07
            High quality sequence stop: 705.
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6086238"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
```

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 38.7%; Score 811.4; DB 5; Length 895;
Best Local Similarity 98.2%; Pred. No. 5e-155;
Matches 863; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

```
Qy        309 AAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTA 368
              || |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 AACCTGGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTA 63

Qy        369 CTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACT 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACT 123

Qy        429 GGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGA 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGA 183

Qy        489 AAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGA 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGA 243

Qy        549 CGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGA 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 CGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGA 303

Qy        609 GTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCG 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCCG 363

Qy        669 AAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCA 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCA 423

Qy        729 GAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAG 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAG 483

Qy        789 GCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCA 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATCA 543

Qy        849 CTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGA 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGA 603

Qy        909 TGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTC 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTC 663

Qy        969 TACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAG 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 TACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTAG 723

Qy       1029 GAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATG-CCCCCCGCTCCTTCAGGCA 1087
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        724 GAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCCCGCTCCTTCAGGCA 783

Qy       1088 GAGCATATACCAGTCCCTTGATCGATATGTTTAAT-AACCCAGCCACGGCTG-CCCCGAA 1145
              ||||||||||||||||||||||||||||||||||| |||||||||||||||| |||||||
Db        784 GAGCATATACCAGTCCCTTGATCGATATGTTTAATAAACCCAGCCACGGCTGCCCCCGAA 843

Qy       1146 TTCACAAAGGGTAAA--TAATTCAACAGGTACTTCCGAA 1182
              |||| ||||| |||   ||| ||  |||||||| ||||
Db        844 ATCACCAAGGGGAAAATAAATCCACAAGGTACTTTCGAA 882
```

RESULT 8
CA450094/c
LOCUS CA450094 814 bp mRNA linear EST 08-NOV-2002
DEFINITION UI-CF-FN0-afb-o-02-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone UI-CF-FN0-afb-o-02-0-UI 3', mRNA sequence.
ACCESSION CA450094
VERSION CA450094.1 GI:24814514
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 814)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FN0-afb-o-02-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FN0"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-FN0 is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h
TAG_LIB=UI-CF-FN0
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 38.3%; Score 803.4; DB 6; Length 814;
Best Local Similarity 99.1%; Pred. No. 2.1e-153;
Matches 807; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1239 GCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTT 1298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         814 GCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTN 755

Qy       1299 TGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGG 1358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         754 TGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGG 695

Qy       1359 AGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGA 1418
```

```
Db        694 AGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGA 635

Qy       1419 AGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCAT 1478

Db        634 AGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCAT 575

Qy       1479 CAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGA 1538

Db        574 CAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGA 515

Qy       1539 ATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCAC 1598

Db        514 ATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCAC 455

Qy       1599 CTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGC 1658

Db        454 CTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGC 395

Qy       1659 AAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGT 1718

Db        394 AAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGT 335

Qy       1719 GACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGC 1778

Db        334 GACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGC 275

Qy       1779 CTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCG 1838

Db        274 CTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCG 215

Qy       1839 GGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTC 1898

Db        214 GGAAGCTTCTGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTC 155

Qy       1899 TGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAAT 1958

Db        154 TGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAAT 95

Qy       1959 GACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGG 2018

Db          94 GACTAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGG 35

Qy       2019 TTTCAATCTTTAAAAAAAAAAAAAAAAAAAAA 2052

Db          34 TTTCAATCTTTACTGTAAAAAAAAAAAAAAAAA 1


RESULT 9
BU184917
LOCUS       BU184917                877 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7825544 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6093473
            5', mRNA sequence.
ACCESSION   BU184917
VERSION     BU184917.1  GI:22698901
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13362  row: 1  column: 18
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1. .877
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6093473"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             37.9%;  Score 794.8;  DB 5;  Length 877;
  Best Local Similarity   97.8%;  Pred. No. 1.2e-151;
  Matches  859;  Conservative    0;  Mismatches   12;  Indels    7;  Gaps    5;

Qy       1118 TTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTT 1177

Db          1 TTAATAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTT 60

Qy       1178 CCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGA 1237

Db         61 CCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGA 120

Qy       1238 AGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCT 1297

Db         121 AGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCT 180

Qy       1298 TTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATG 1357

Db         181 TTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATG 240

Qy       1358 GAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGG 1417

Db         241 GAGAACACGACGTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGG 300

Qy       1418 AAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCA 1477

Db         301 AAGAAAATGAGACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCA 360

Qy       1478 TCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGG 1537

Db         361 TCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGG 420

Qy       1538 AATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCA 1597

Db         421 AATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCA 480

Qy       1598 CCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTG 1657

Db         481 CCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTG 540

Qy       1658 CAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTG 1717

Db         541 CAAAGCCGCCTTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTG 600

Qy       1718 TGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGG 1777

Db         601 TGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGG 660

Qy       1778 CCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTC 1837

Db         661 CCTCTCCGGTTCTCCCTTGCGGAATGATGGGCGCATCCTGTCTGCCA-GTGCTGACGGTC 719

Qy       1838 GGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCT 1897

Db         720 GGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTGCTTAAGC-AATCATGCTTCT 778
```

```
Qy       1898 CTGTTTCAC--GTAGTTGGGTTGACAAGTTT--CTGCCTTTAAGATAAATGAGT-AATAG 1952
             ||||||| |    | | |||||||| |||||     | ||||||||||||||| |||||
Db        779 CTGTTTCCCCGTAATTGGGGTTGACCAGTTTTCTGCCTTTTAAGATAAATGAGTAAATAG 838

Qy       1953 TCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCC 1990
             |||||||||||||||||||||||||||||||||| | ||
Db        839 TCTAATGACCAGCTCAGCCATTTAAAATATTTTCCTTC 876


RESULT 10
BQ688252
LOCUS       BQ688252                916 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8034808 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6207009
            5', mRNA sequence.
ACCESSION   BQ688252
VERSION     BQ688252.1  GI:21813568
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 916)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2362  row: g  column: 10
            High quality sequence stop: 673.
FEATURES             Location/Qualifiers
     source          1. .916
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6207009"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_110"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             37.8%;  Score 792.2;  DB 5;  Length 916;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-151;
  Matches  794;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        308 TAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCT 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCT 60

Qy        368 ACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAAC 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAAC 120

Qy        428 TGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATG 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGGACATGTCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATG 180

Qy        488 AAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTG 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTG 240

Qy        548 ACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAG 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAG 300

Qy        608 AGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCC 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCC 360

Qy        668 GAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTC 727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAACGCACTCAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTC 420

Qy        728 AGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGA 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGA 480

Qy        788 GGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATC 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACATTATTATC 540

Qy        848 ACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTG 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTG 600

Qy        908 ATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCT 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCT 660

Qy        968 CTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTA 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAATGTGGTTA 720

Qy       1028 GGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCA 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCA 780

Qy       1088 GAGCATATACCAGTCCC 1104
              |||||||||||   ||
Db        781 GAGCATATACCAAGTCC 797


RESULT 11
BI333892
LOCUS       BI333892                812 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION  602999431F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141574 5',
            mRNA sequence.
ACCESSION   BI333892
VERSION     BI333892.1  GI:15018549
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 812)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11348  row: n  column: 07
            High quality sequence stop: 812.
```

```
FEATURES             Location/Qualifiers
     source          1. .812
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5141574"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN

  Query Match             37.0%;  Score 775.2;  DB 3;  Length 812;
  Best Local Similarity   99.4%;  Pred. No. 1.2e-147;
  Matches  799;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy        709 GAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAG 60

Qy        769 GCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCA 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGCA 120

Qy        829 AACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGG 180

Qy        889 TGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAA 240

Qy        949 GAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATC 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGATC 300

Qy       1009 GAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATG 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATG 360

Qy       1069 CCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCA 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCA 420

Qy       1129 GCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCC 1188
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        421 GCCACGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCTGAAGATCCC 480

Qy       1189 AGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTG 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTG 540

Qy       1249 AAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGA 600

Qy       1309 GATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGAC 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGAC 660

Qy       1369 GTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAG 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAG 720

Qy       1429 ACAGAAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCAT-CAGCACCGT 1487
              |||||||||||||||||| |||||||||| |||||||||||||||||||| ||||||||
Db        721 ACAGAAGCAGTGACCGTG-CCACGCCAAGACCCACACCAGTGAGAAGCATCCAGCACCGT 779

Qy       1488 GAACTTGTCTGAGAATAGCAGTGT 1511
              |||||||||||||||||||||| ||
```

```
Db        780 GAACTTGTCTGAGAATAGCAGGGT 803
```

```
RESULT 12
BM981661/c
LOCUS       BM981661                779 bp    mRNA    linear   EST 21-FEB-2003
DEFINITION  UI-CF-EN1-adi-m-10-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-adi-m-10-0-UI 3', mRNA sequence.
ACCESSION   BM981661
VERSION     BM981661.1  GI:19604378
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 779)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
             The following repetitive elements were found in this cDNA
            sequence: 1-21, >AT_rich#Low_complexity
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .779
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EN1-adi-m-10-0-UI"
                     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                     Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EN1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EN1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Cystic Fibrosis
                     Epithelial Cells. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CTGCTCAGGT.
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-EN1
                     TAG_SEQ=CTGCTCAGGT"
ORIGIN

  Query Match             36.7%;  Score 768.4;  DB 3;  Length 779;
  Best Local Similarity   99.1%;  Pred. No. 3e-146;
```

```
  Matches  772;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1269 TGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCAT 1328
             ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        779 TGCGGGCTCCAACAGACCCTTACTCAGCTTGCNACAGGGAGATGTCATCACGCTGCTCAT 720

Qy       1329 CCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTG 1388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 CCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTG 660

Qy       1389 GTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCC 1448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCC 600

Qy       1449 CACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAG 1508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 CACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAG 540

Qy       1509 TGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAG 1568
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        539 TGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAG 480

Qy       1569 GAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGAC 1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 GAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGAC 420

Qy       1629 CGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCC 1688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 CGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGCGGAGAAAACCC 360

Qy       1689 CTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCG 1748
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 CTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCG 300

Qy       1749 ATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGG 1808
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGG 240

Qy       1809 CGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAA 1868
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAA 180

Qy       1869 TGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCT 1928
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 TGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCT 120

Qy       1929 GCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTT 1988
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         119 GCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTT 60

Qy       1989 CCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAA 2047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          59 CCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAA 1

RESULT 13
BU500695
LOCUS       BU500695                943 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_7860835 NIH_MGC_64 Homo sapiens cDNA clone IMAGE:6109543
            5', mRNA sequence.
ACCESSION   BU500695
VERSION     BU500695.1  GI:22802021
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: NIMH/LOG
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2355  row: j  column: 08
            High quality sequence start: 43
            High quality sequence stop: 453.
FEATURES             Location/Qualifiers
     source          1. .943
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6109543"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_64"
                     /note="Vector: pOTB7a; Site_1: CeuI; Site_2: SceI; This
                     library is a size selection of NIH_MGC_35, from 3.0-4.5
                     kb. Size selection done at the National Institute of
                     Mental Health, NIH. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           36.4%;  Score 762.6;  DB 5;  Length 943;
  Best Local Similarity   95.3%;  Pred. No. 4.7e-145;
  Matches  865;  Conservative    0;  Mismatches   32;  Indels   11;  Gaps    8;

Qy       1151 AAAGGGTAAATAATTCAACAGGT-ACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCG 1209
             || ||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||
Db         17 AAGGGGTAAATAATTCAACAGGTAACTTCTGAAGATCCCAGTTTACAGCGATCAGTTTCG 76

Qy       1210 GTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACT 1269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACT 136

Qy       1270 GCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATC 1329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATC 196

Qy       1330 CCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGG 1389
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 CCCGAGGAAAAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGAGGGGTTGG 256

Qy       1390 TTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCC 1449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCC 316

Qy       1450 ACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGT 1509
             |||||||||||||||||||||||| |||||||||||||||||||||| |||||||||
Db        317 ACGCCAAGCCCCACACCAGTGAGAAACATCAGCACCGTGAACTTGTCTGAAAATAGCAGT 376

Qy       1510 GTTGTCA-TCCCCCCACCCGACTACTTGGAATGCTTGT---CCATGGGGGCAGCTGCCGA 1565
             ||||||| |||||||||||||||||||| | |   | |   | |||||||||||||||||
Db        377 GTTGTCATTCCCCCCACCCGACTACTTTGGAATGCTTTGTCCATTGGGGGCAGCTGCCGA 436

Qy       1566 CAGGAGAGCAGATTCGGCCAGGACGACA-TCCACCTTTAAGGCCCCAGCGTCCAAGCCCG 1624
             ||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        437 CAGGAGAACAGATTCGGCCAGGACGACATTCCACCTTTAAGGCCCCAGCGTCCAAGCCCG 496

Qy       1625 AG-ACCGCGGCTCCTAACGATGCCAACGGGA--CTGCAAAGCCGCCTTTTCTCAGCGGAG 1681
             || |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        497 AGAACCGCGGCTCCTAACGATGCCAACGGGAACTTGCAAAGCCGCCTTTTCTCAGCGGAG 556

Qy       1682 AAAACCCCTTTGCCACTGTGAAAC-TCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 1740
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        557 AAAACCCCTTTGCCACTGTGAAACTTCCGCCCGACTGTGACGAATGATCGCTCGGCACCC 616

Qy       1741 A-TCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGG 1799
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 ATTCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGTTCTCCCTTGCGG 676
```

```
Qy       1800 AATGATGGGCGCATCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCC 1859
             ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 AATGATGGGCGCGTCCTGTCTGCCACGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCC 736

Qy       1860 TAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGA 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 TAACTCTAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGA 796

Qy       1920 CAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAA 1979
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        797 CAAGTTTCTGCCTTTAAAATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAA 856

Qy       1980 TATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAA 2039
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        857 TATTTTCTTCCTATTCTGTTCAAAAAACAGTAAACTTGGTTTCAATCTTTANNANNNNNN 916

Qy       2040 AAAAAAAA 2047
                  |||
Db        917 NNNNNAAA 924
```

RESULT 14
BI335025

LOCUS	BI335025 996 bp mRNA linear EST 30-JUL-2001
DEFINITION	602999031F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141190 5', mRNA sequence.
ACCESSION	BI335025
VERSION	BI335025.1 GI:15019682
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 996)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11347 row: n column: 07 High quality sequence stop: 827.

FEATURES	Location/Qualifiers
source	1. .996 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5141190" /tissue_type="cervical carcinoma cell line" /lab_host="DH10B" /clone_lib="NIH_MGC_12" /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN

Query Match 36.3%; Score 760.4; DB 3; Length 996;
Best Local Similarity 96.5%; Pred. No. 1.3e-144;
Matches 853; Conservative 0; Mismatches 21; Indels 10; Gaps 7;

```
Qy        708 GGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGA 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGA 60

Qy        768 GGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 120

Qy        828 AAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCG 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCG 180

Qy        888 GTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGA 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGA 240

Qy        948 AGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGAT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AGAAATAAAGACCCCAGCCTCTACCCCCGTGTCTGGAACTCCTCAGGCTTCACCCATGAT 300

Qy        1008 CGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGAT 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CGAGAGAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGAT 360

Qy        1068 GCCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCC 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GCCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCC 420

Qy        1128 AGCCACGGCTG-CCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATC 1186
              ||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db         421 AGCCACGGCTGCCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCTGAAGATC 480

Qy        1187 CCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAG 1246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAG 540

Qy        1247 TGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGG 1306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGG 600

Qy        1307 GAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACG 1366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACG 660

Qy        1367 ACGTGTCCAAGGCGAGGGGTTGGTT-CCCGTCGTCGTACACGAAGTTGCTGGAAGAAAAT 1425
              |||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||||
Db         661 ACGTGTCCAAGGCGAGGGGTTGGTTCCCCGTCGTCGTACACGAAGTTGCTGGACGAAAAT 720

Qy        1426 GAGA-CAGAAGCAGTGACCGTGCCCACG--CCAAGCCCCACACCAGTGAGAAGCATCAGC 1482
              |||| |||||||||||||||||||||  | | |||||||||||||||||||||||||||||
Db         721 GAGACCAGAAGCAGTGACCGTGCCCACGGCCAAGGCCCCACACCAGTGAGAAGCATCAGC 780

Qy        1483 ACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCCACCCGACTACTTGGAATGC 1542
              |||||||||||||||||||||||||||||||| |      |||||||      | ||||||
Db         781 ACCGTGAACTTGTCTGAGAATAGCAGTTGTTGTCATCCCCCACC---GATACTTGAATGC 837

Qy        1543 TTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAG 1586
              |||| |||||||||||| |||||||||||| |||| |||||||| |
Db         838 TTGT-CATGGGGGCAGTTGCCGACAGGA-AGCAAATTCGGCCCG 879
```

RESULT 15
BG761459

LOCUS	BG761459 933 bp mRNA linear EST 15-MAY-2001
DEFINITION	602718666F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858696 5', mRNA sequence.
ACCESSION	BG761459
VERSION	BG761459.1 GI:14072112
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 933)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1713  row: c  column: 17
            High quality sequence stop: 851.
FEATURES              Location/Qualifiers
     source           1. .933
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4858696"
                      /tissue_type="melanotic melanoma, high MDR (cell line)"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_49"
                      /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Size-selected >500bp for average insert size
                      1.8kb. Library constructed by Ling Hong in the laboratory
                      of Gerald M. Rubin (University of California, Berkeley)
                      using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                      II RT (Life Technologies). Note: this is a NIH_MGC
                      Library. |"
ORIGIN

  Query Match           36.1%;  Score 757.2;  DB 2;  Length 933;
  Best Local Similarity   95.8%;  Pred. No. 6e-144;
  Matches  885;  Conservative    0;  Mismatches   28;  Indels   11;  Gaps   10;

Qy       1013 GAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCC 1072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GAAGCAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCC 61

Qy       1073 CCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCA 1132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCA 121

Qy       1133 CGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTT 1192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CGGCTGCCCCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTT 181

Qy       1193 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA 1252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA 241

Qy       1253 CCATCTTCCCGC-ACACTGCGGGCTCCAA-CAAGACCTTACTCAGCTTTGCACAGGGAGA 1310
             |||||||||||| |||||||||||||||| |||||||||||||||||||||||||||||
Db        242 CCATCTTCCCGCGACACTGCGGGCTCCAAGCAAGACCTTACTCAGCTTTGCACAGGGAGA 301

Qy       1311 TGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTATGGAGAACACGACGT 1370
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        302 TGTCATCACGCTGCTCATCCCCGAGGAGACGGATGGCTGGCTCTATGGAGAACACGACGT 361

Qy       1371 GTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGAC 1430
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GTCCAAGGCGAGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGAAGAAAATGAGAC 421

Qy       1431 AGAAGCAGTGACC-GTGCCCACGCCAAGCCCCA-CACCAGTGAGAAGCATCAGCACCGTG 1488
             |||||||||||||| |||||||||||||||||| |||||||||||||||||||||||||
Db        422 AGAAGCAGTGACCTGTGCCCACGCCAAGCCCCAGCACCAGTGAGAAGCATCAGCACCGTG 481

Qy       1489 AACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCC 1548
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        482 AACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTATCC 541

Qy       1549 ATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        542 ATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCC 601

Qy       1609 CCAGCGTCCAAG-CCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCC 1667
             |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CCAGCGTCCAAGTCCCGAGACCGCGGCTCCTAACGATGCCAACGGGACTGCAAAGCCGCC 661

Qy       1668 TTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGA 1727
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        662 TTTTCTCAGCGGAGAAAACCCCTTTGCCACTGTGAAACTCCGTCCGACTGTGACGAATGA 721

Qy       1728 TCGCTCGGCACCCATCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGT 1787
             |||||| |||||||||||||||||||||||| | ||||||||||||||    | | |||
Db        722 TCGCTCAGCACCCATCATTCGATGAGAGGACAGGC-AGGACTCTCCCGGGCTCTCCGGTT 780

Qy       1788 TCTCCCTTGCGGAATGATGGGCGCATCCTGTC-TGCCACGTGCTGACGGTCGGGAAGCTT 1846
                || |  |||||||||||||| ||||||| |||||||||||||||||||||||||| |
Db        781 CTCCCTTGCGGGAATGATGGGCGCGTCCTGTCTTGCCACGTGCTGACGGTCGGGAAGCCT 840

Qy       1847 CAGTGGA--GAGGCCTAACTCTAATGTC-GCCTGCTTAAGCAAAT-CATGCTTCTCTGTT 1902
             |||||||    |||||||||||| ||||| |||||||||||||||| ||||||||||||||
Db        841 CAGTGGAAGAGGGCCTAACTCTTATGTCGGCCTGCTTAAGCAAATCCATGCTTCTCTGTT 900

Qy       1903 TCACGTAGTTGGGTTGACAAGTTT 1926
             || |||||||||||| ||  | |
Db        901 TCCCGTAGTTGGGTTGCCAGTTCT 924


Search completed: March 20, 2006, 15:59:48
Job time : 8390 secs
```